D

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

VOLUME XIX, 1967

PUBLISHED FOR
The Experimental Psychology Society
BY
W. Heffer & Sons Ltd. Cambridge

The Quarterly Journal of Experimental Psychology

Volume XIX

CONTENTS

SHORTER ARTICLES AND NOTES

APPARATUS

REVUE CANADIENNE CANADIAN JOURNAL DE PSYCHOLOGIE OF PSYCHOLOGY

ANNOUNCEMENT

The Journal begins bi-monthly publication in 1967. Issues will appear in February, April, June, August, October and December. This change has been made necessary by an increased number of acceptable manuscripts being submitted, and the decision by the Canadian Psychological Association to keep the average publication lag to the nine months it has been over the last several years.

The annual subscription rate will be increased from \$6.00 to \$10.00, and the price of individual issues from \$1.75 to \$2.00.

Concessionary subscription rate to members of the Experimental Psychology Society \$8.00.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIX 1967 Part 1

SELECTIVE ATTENTION: PERCEPTION OR RESPONSE?

BY

ANNE TREISMAN and GINA GEFFEN

From the M.R.C. Psycholinguistics Research Unit, Oxford

Does our limited capacity in selective listening tasks arise primarily in perception or in response organization? To examine this, subjects were given two dichotic messages, one primary and one secondary, and had to make two different responses: the primary response was to "shadow" the primary message; the secondary response was to tap on hearing certain target words in either message. Since the secondary response was identical for the two messages, any difference in its efficiency with the two messages must be due to a failure in perception of the secondary message. Any interference between the primary and secondary responses (repeating and tapping) to target words in the primary message must be due to a limit in performing simultaneous responses, since if either was correctly performed the target word must have been perceived. The results clearly showed that the main limit is perceptual.

Various target words were used to investigate the nature of the perceptual and response limits. Factors investigated were (1) the information content of the target words, (2) their range of meanings, (3) their grammatical class, and (4) the compatibility between stimuli and responses. A relative lack of response competition was found, which might be due to successive organization of the two responses at different stages in the perceptual sequence. The results were interpreted in terms of signal detection theory and the effects of reduced signal-to-noise ratio produced by inattention were compared with those produced by an external masking noise.

INTRODUCTION

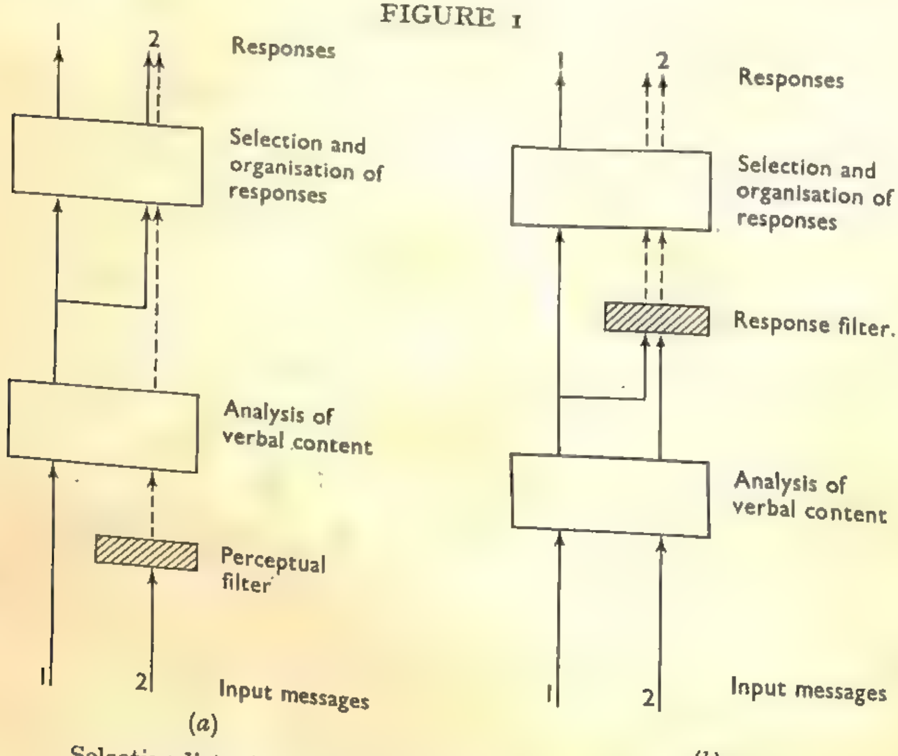
The characteristics of human selective attention have recently been explored in some detail, typically by experiments requiring subjects to respond selectively to one of two or more simultaneous speech messages. When the two messages come from different sources, subjects can repeat one back very efficiently, but can usually report nothing of the verbal content of the other (Cherry, 1953), apart from a few highly important or relevant words (Moray, 1959, Treisman, 1960). If we ask the subject specifically to recall single target words presented to one ear, his ability to repeat the words on the other ear is totally disrupted at the times when the target words occur (Mowbray, 1964). This limit to performance is clearly not due to a shortage of ears or mouths, since either message can be clearly heard, and since the verbal responses required are successive not simultaneous. But we can still ask whether the limit to our capacity for perceiving speech arises on the perceptual or on the response side of the brain's central communication channel. Can we only analyse and identify half the incoming words or can we only organise memory storage and response for one of the two messages?

Broadbent (1958) favours the perceptual hypothesis and suggests that a "selective filter" in the brain rejects the unwanted message before its content is fully analysed. Deutsch and Deutsch (1963) prefer the explanation that all stimulus inputs are fully analysed and that selection is made only to determine responses and memory.

Reynolds (1964) gives a similar account in terms of competition between responses and assumes that the second message in Cherry's type of experiment is "not a potent elicitor of responses" except when it contains the subject's own name or other highly relevant signals. One obvious way of separating response from stimulus competition is to ask for two different responses from one and the same stimulus and to compare this with the same two responses made to two different simultaneous stimuli. The experiment we shall describe was an attempt to throw more light on the nature of the limit to human speech transmission and on the level at which it occurs.

To test how far attention is a feature of perception rather than of response, we can compare the same response made to an attended and an unattended message. To test how far the limit affects the performance of responses, we see what interference a second response to the same stimulus causes in the performance of a primary response. We can combine these two problems into one experimental test by presenting two messages and requiring two responses, one of each being given priority by the instructions. The primary message and response are chosen (on the basis of previous findings) to occupy most of the limited capacity available to the subject. The primary response is made to the primary message only and the other response to both messages. Both are made immediately the stimulus is given, so no memory limits are involved in determining performance. The two hypotheses are illustrated in Figure 1.

FIGURE 1



Selective listening (*a*) assuming a limit to perceptual capacity and (*b*) assuming a limit to response capacity.

Figure 1*a* assumes a perceptual limit, with a "filter" reducing the perceptual analysis of message 2, but no limit to the responses which can be made to perceived signals. Response 2 is therefore made much more efficiently to message 1 than to message 2. Figure 1*b* assumes that the two messages are perceived equally well, but

that only one response can be efficiently performed. Response 2 is therefore equally inefficient to message 1 and to message 2. In this experiment, the two responses to the primary message are made to the same stimulus, which must have been correctly perceived if either response is made correctly. Any reduction in the efficiency of the primary response produced by the second response to the primary message must therefore be due to response rather than stimulus competition. In the experiment to be reported, the two messages were prose passages, played one to each ear of the subject over headphones. The subject attended to one of the two (the primary message) and repeated it back continuously as he listened to it (the primary response). The secondary response, which was made to both messages, was to tap with a ruler whenever a particular target word was heard in either message.

The main aim was to compare perceptual and response competition in the selective listening task, but it is of interest also to explore the nature of the limit in more detail: (1) What effect will verbal characteristics of the target words have on performance? Since Broadbent put forward his filter theory of selective attention in 1958, it has generally been assumed that the main limit in selective attention is determined by the information content of the messages. We can compare three ways of varying the information content of the target words: (*a*) varying their transition probability (this was done by fitting the target words into the verbal context of the passage or inserting them at random points); (*b*) varying the number of target words by using members of large or small ensembles (for example comparing "any digit" with the single word "Boat"); (*c*) increasing the linguistic or semantic ambiguity without changing the number of phonetic patterns by using words with several meanings (e.g. "Fit") or several homophones (e.g. "Right, Rite, Write, Wright"). These variables might have different effects on perceptual and on response competition. If the hypothesis of limited *perceptual* capacity is correct, some clear differences should also emerge between target words in the primary and secondary message with respect to these linguistic variables. For instance if the secondary message is filtered out before analysis of its verbal content, we should not expect changes in context or meaning to affect performance. The subject would react to the target words, if at all, simply as particular speech sounds.

(2) If the responses compete, what is the nature of the competition? If two responses are made to the same stimulus (for example the verbal response of repeating the target word and the manual response of tapping to it), are they determined at the same stage of perceptual analysis, or is the manual response triggered simply by recognition of the speech sound before its syntactic and semantic role in the sentence is analysed? When one states that a word has been "perceived" one is not making an all-or-nothing assertion: it may be that no stimulus reaching our nervous system is ever fully classified in all possible ways. We probably analyse chiefly those features relevant to the particular response we wish to make. If we assume that speech perception is a hierarchical process, in which categorizations may be made at a number of different levels, such as the physical sound, the phonemic pattern, the word, the syntactic structure and the semantic interpretation, it may be possible for different responses to be selected and programmed at different stages in the sequence rather than all being dependent on its completion. In the present experiment the particular acoustic pattern of the target words could have been sufficient signal for the tapping response, but the repeating response probably required a higher level of analysis, since subjects were repeating the whole passage rather than isolated words. Other experiments have shown that they can only do this efficiently, at the speed we used, when they make use of the general contextual redundancy, implying some recognition of syntax and meaning. If this redundancy is reduced, the repeating

response breaks down (Moray and Taylor, 1958). If the tapping and repeating responses do not interfere, this might be due to their being "cued" at different points in the perceptual sequence. This suggestion was not tested directly, but the results give some indications which will be examined in the discussion.

(3) We investigated the effect of stimulus-response compatibility on attention. Can one, by choosing a response which is closely related to the stimulus, bypass the usual limited capacity decision channel? In reaction time tasks there is now considerable evidence (Leonard, 1961; Mowbray, 1960; Davis, Moray and Treisman, 1961; Broadbent and Gregory, 1962) that the more compatible, overlearned, natural and automatic the relation between stimulus and response, the less effect is produced by increases in information content. The subject appears to function as a multichannel system in which decisions are taken in parallel rather than sequentially. Would the same be true of selective listening tasks? Moray and Jordan (1966) suggest that compatibility may be equally important here.

(4) Finally we hoped to compare the effects of auditory noise and of inattention on the perception of speech. Broadbent and Gregory (1963) and Treisman (1960, 1964) suggested that the perceptual filter mediating selective attention might "attenuate" unwanted messages rather than block them completely. If the effect of inattention is to reduce the signal-to-noise ratio of all but the selected message, one might expect the resultant behaviour to resemble that produced by an external masking noise.

Method

Apparatus and stimulus materials. Two Ferrograph twin-track tape-recorders were used. The experimental messages were presented on one tape-recorder and the responses were recorded on one track of the second while the primary message was re-recorded on the other track. The messages were presented dichotically to subjects through a pair of Brown moving-coil headphones, and both responses were recorded through a microphone. The prose passages were all extracts from "Lord Jim" by Conrad, some of them modified slightly to allow target words to be inserted in context. Both were recorded by the same woman speaker. The primary message started two or three words before the secondary one and they finished together; each was 150 words long and lasted about 1 min. Three target words were inserted in each passage of every pair at random points, with the restriction that none occurred in the first or last 10 words or within less than eight words of another target word in either the same message or the competing message on the other ear. The intensities of these target words were measured using a Marconi valve voltmeter, and were later correlated with the subjects' performance.

Four different tape-recordings were used. The first three were essentially similar, in that each used the same five classes of target words and they were designed to test the same theoretical points. However each recording used a different set of prose passages and different examples of each class of target word, in order to control for accidental differences in difficulty of particular words or passages. The target words used in these three tapes are given in Table I, together with descriptions of the variables being investigated. Each type of target word was given in context in two passages (e.g. "her big

TABLE I

Target Words Used to Test the Main Experiment Variables

Variable	*High stimulus information*	*Adjective, one main meaning*	*Noun, one main meaning*	*Several meanings*	*High stimulus-response compatibility*
Group A	Any Digit	Tall	Boat	Right	Tap
Group B	Any Colour	Hot	Trees	Fit	Tap
Group C	Any Part of the face	Tired	Night	Point	Tap

clear *eyes* would remain fastened on us . . .") and out of context in two other passages (e.g. "waste ground interspersed *cheek* with small patches of . . ."). In each condition (e.g. "Hot" in context) there were therefore six target words in the primary messages (three in each of two passages) and six in the corresponding secondary messages. These recordings were used to investigate the following variables, and to compare their effects on perception of and response to the target words: (1) transition probabilities from the verbal context; (2) ensemble size of the target words; (3) variety of possible meanings of target words all sharing the same phonetic form; (4) different grammatical forms of the target words; (5) stimulus-response compatibility.

The fourth tape-recording included the following target words:

(i) "From" and "But," each in and out of context, to compare the efficiency of responses to functional, non-lexical words with those to the nouns and adjectives of the first three recordings.

(ii) "Right" in context; homophones of "Right" (i.e. "Write," "Rite" and "Wright") in context; "Right" in the primary message with homophones of "Right" in the secondary message. There are two main questions: firstly would the homophones be any more difficult than the single word, that is would they function as several different target words in the same way as the digits, colours or parts of the face, or could they all be treated as a single, target, speech sound. Secondly, would subjects find it difficult to avoid tapping to homophones when instructed to tap only to "Right," indicating that the tapping response was initiated at some stage before the meaning was analysed? Both answers should help to locate the stage in perceptual analysis at which the tapping response was initiated, and to show whether this differed for attended and unattended message.

(iii) "Hot" in context in two primary messages paired with "Hot" out of context in the secondary messages, and two pairs with the reverse arrangement. In all other conditions the target words were either in context on both ears or out of context on both. The context of the primary message could therefore conceivably facilitate perception of target words in the secondary message as well. For example when the target word was "Hot," the passage might be about a parched and sweating man under a blazing sun in the desert. This restricted subject-matter could generally lower thresholds for perception of "Hot" in either passage. Pairing target words in context on one ear with target words out of context on the other controls for this possibility and ensures that any effect of context is restricted to the passage in which the target word itself occurs in context.

General procedure. All subjects except the control groups were treated as follows. They were given some practice trials at repeating back one of two simultaneous speech passages, until they were doing this fluently, and also some practice at tapping to target words (not those used in the experiment proper). They were then given the experimental passages with the primary one always on the right ear, and were asked to repeat back this passage, keeping their attention fixed on the right ear. They were told before each pair of passages what the target word would be (e.g. "Hot in context" or "any colour out of context") and they were asked to tap if ever they heard this target word in either ear; they were not to shift their attention to the secondary message, since we were interested in seeing whether they heard it *despite* the fact that they were attending to something different. It was emphasized that if they shifted their attention, they would miss some words of the primary passage and so fail in the primary task. They were asked, after each passage in which they tapped to a word in the secondary message, whether they felt they had shifted their attention in order to hear the word or whether it had "just come through" while they were attending to the right ear message. They seemed quite able to distinguish these two cases, as shown in the results. Cases in which they had shifted before hearing the target word were not included in the results, since the secondary message had effectively become the primary message for those few moments.

The first passages containing each type of target word were presented in different random orders, and the second passages in the reverse order, to counterbalance any effects of practice or fatigue. The experiments lasted about 1½ hr. with a 10-min. break in the middle. Subjects were questioned about which conditions seemed most difficult, and were also asked whether the loudness of the two passages remained approximately equal throughout.

Further differences in the procedure adopted for control groups are given in the following section on subjects and design.

Subjects and design of experiments. The volunteer subjects were undergraduates at Oxford University (none reading Psychology); they were paid four shillings an hour. Their hearing was approximately equal in both ears and each subject equated the loudness of the tape-recordings in the two ears for himself before starting the experiment. They were divided into a number of different groups.

Group A, consisting of 42 subjects, was used in Experiment 1, to investigate the variables summarized in Table I. Fourteen subjects were tested with each of the first three tape-recordings, following the general procedure described above. Since they all had essentially the same conditions, their results were analysed together.

Group B, consisting of nine subjects, was used in the masking Experiment, 2. Each of the same Group A passages was presented singly, masked by noise, and the subjects were asked simply to tap whenever they heard one of the target words. Thus their attention was focussed on the single message and the single response of tapping to target words. Each primary and secondary message in each of the first three tape-recordings was heard by three of the subjects in Group B. The signal-to-noise ratio was adjusted in a pilot experiment to give approximately 50 per cent. correct responses.

Group C, comprising 11 new subjects, was used in Experiment 3, in order to clarify some points arising from Experiment 1. They followed the same procedure as Group A but were given the new target words and passages of the fourth tape-recording.

Finally Group D, a further 10 subjects, was used in Experiment 4 to check on the effect of cerebral dominance. Since all other subjects attended to the right ear and tapped with the right hand, there might have been some cerebral asymmetry favouring the primary message. Group D followed the same procedure as Groups A and C, and heard the third tape-recording from Experiment 1, but one primary message containing each type of target word was played to the right ear and one to the left. Each passage of a pair was repeated by half the subjects when on the right ear and by half the subjects when on the left. All target words on the right ear were tapped to by the right hand and all target words on the left ear by the left hand.

Results

Correct responses. The recorded responses were analysed as follows: counts were made of the target words correctly tapped to, the number of target words receiving both tapping and repeating responses, the number of target words receiving only one of the two responses and the number receiving neither response. These were converted to percentages and are given in Tables II and III. Analyses of variance were

TABLE II

Percentage of Target Words Receiving Verbal or Manual Response

Target word		Digits, Colours, Parts of face		Right, Fit, Point		Tall, Hot, Tired		Boat, Trees, Night		But, From (Group C)		Tap	Mean	
In or out of context		In	Out	In	Out	In	Out	In	Out	In	Out	Out	In	Out
Primary message	Repeat	92·9	55·2	97·0	68·5	95·9	68·1	98·1	68·5	95·5	75·0	68·5	95·9	67·1
	Tap	88·9	67·4	93·3	84·4	95·6	85·2	95·1	90·7	75·8	81·9	93·3	89·7	81·9
Secondary message	Tap	8·5	2·6	3·7	7·8	15·6	8·1	14·8	9·9	4·5	6·1	7·8	9·4	6·9

All results are from Groups A, B and C except for target words "But" and "From" which are results from Group C.

carried out on the number of correct tapping responses for all subjects in Group A together. One analysis was made on all conditions except the nouns and the target word "Tap." The other analysis was made just on the "Tap" and "Tall, Hot or Tired" out of context conditions (since "Tap" was never given in context). The main points which emerged are as follows:

(1) An overwhelming majority of tapping responses was made to the primary message rather than the secondary one, a mean of 86·5 per cent. compared with 8·1 per cent. This strongly indicates a perceptual limit in selective listening.

TABLE III

Percentage of Target Words in Primary Message Receiving Both, One or Neither Response

Target words,	*Digits, Colours, Parts of face*		*Right, Fit, Point*		*Tall, Hot, Tired*		*Boat Trees, Night*		*But, From, (Group C)*		*Tap*
In or out of context	*In*	*Out*	*In*	*Out*	*In*	*Out*	*In*	*Out*	*In*	*Out*	*Out*
Both responses	87·3	51·1	92·1	66·6	94·4	65·9	95·1	67·9	75·8	69·0	67·8
One response	7·1	20·4	6·0	19·7	2·6	21·5	3·1	23·5	19·6	18·8	26·3
Neither response ..	5·6	28·5	1·9	13·7	3·0	12·6	1·8	8·6	4·6	12·2	5·9

(2) The target words in the primary message which fitted into the verbal context received more responses than those occurring at random points, ($p < 0{\cdot}001$, $V.R. = 31{\cdot}7$, $d.f.$ 1,41). In the secondary message the difference was also significant, though slighter ($p < 0{\cdot}025$, $V.R. = 6{\cdot}5$, $d.f.$ 1,41). Here it was due mainly to the specific, lexical words, and was in fact reversed for the words of many meanings. A possible explanation here is that the facilitation was due to the context of the primary message. This would explain why context favoured only the specific lexical words and not the function words or the words of many meanings, since these did not restrict the general theme of the primary message. When target words in context on one ear were paired with target words out of context on the other ear, the effect of context on the secondary message disappeared. Group C made 14 per cent. correct tapping responses to "Hot" out of context and 15 per cent. to "Hot" in context. This confirms that part at least of the facilitating effect of verbal context in other conditions was due to the related subject-matter of the primary message. Moreover, a later experiment (not yet published) using the same task has shown no effect at all of verbal context in the secondary message. This reinforces our belief that the apparent facilitation here is an artefact.

(3) There were some significant differences between the different types of target words in both primary and secondary messages. The classes of words (digits, colours and parts of the face) received significantly fewer responses than the single words, particularly when they were out of context. A Scheffé test showed that the difference was significant, $p < 0{\cdot}001$ for both primary and secondary messages. The words with many meanings were as easy to tap to as the words with one main meaning in the primary message, but they were more difficult in the secondary message. This latter finding is surprising in the light of other results, particularly of Group C's experiment with the homophones of "Right." These are given in Table IV. In

TABLE IV

Per cent Tapping Responses to "Right" and its Homophones

Target word	1. *Right* (*correct*)	2. *Homophones* (*correct*)	3. *Homophones* (*incorrect*)	4. *Right* (*correct in primary message*); *Homophones* (*incorrect in secondary message*)
Primary message	91	94	30	88
Secondary message ..	5	8	3	8

Condition 1, subjects were presented with and asked to tap only to "Right"; in Condition 2, they were presented with any of the four homophones, "Right," "Write," "Rite" and "Wright" and were asked to tap to any of them. This condition proved no more difficult than the first, with either primary or secondary message. In Condition 3, they were asked to tap only to "Right" but were presented only with its homophones, and in Condition 4 they were again asked to tap only to "Right," but while only "Right" was presented in the primary messages only its homophones were included in the secondary messages. Condition 3 shows that subjects could successfully discriminate and avoid about two-thirds of the incorrect homophones in the primary message. However in the secondary message they tapped to as many of the incorrect homophones in Conditions 3 and 4 as they did to the correct words in Conditions 1 and 2. They could not therefore have heard the verbal context which differentiated the four homophones. Given this proof that the homophones were not identified as such, we can see no obvious explanation for the relative difficulty of "Right" and its homophones in the secondary message compared with the single adjectives and nouns.

(4) Differences in grammatical form only affected performance taking the extreme comparison of lexical items versus functional parts of speech The nouns and adjectives gave identical results, but "From" and "But" evoked fewer tapping responses when they were *in* context. However they were *repeated* just as efficiently as the nouns and adjectives.

(5) The target word "Tap" in the primary message received significantly more tapping responses than the adjectives "Tall," "Hot" or "Tired" out of context ($p < 0{\cdot}05$, $V.R. = 4{\cdot}61$, $d.f. = 1, 41$), but in the secondary message the difference disappeared.

We return to these results in the Discussion, where we try to relate them to the problems raised in the Introduction.

Interference. The next result we analysed was the amount of interference with the repeating response caused by the target words in the primary and secondary messages. We compared the interference in those cases where the subject tapped correctly to the target word and in the cases where he failed to tap. For this purpose, the words of the primary message were divided into four categories: (i) the target words in the primary message, or the words in the primary message which coincided with target words in the secondary message; (ii) the three words preceding these; (iii) the five words succeeding them; (iv) all other words. The percentage of these classes of words which showed errors or omissions in repeating was calculated separately for the occasions when subjects tapped and when they failed to tap and separately for target words in the primary and in the secondary message. On analysis of variance, no significant difference in interference emerged due to differences in the target words used, so the results given in Table V are the mean percentages over all the target words.

The table shows that tapping to target words in the primary message interfered slightly with the repeating response both to the target word itself and to the succeeding five words, a mean of about 11 per cent. errors compared with about 7 per cent. to the other words. (The variance due to which words were being repeated—target, three before, five after or other words—was significant, $p < 0{\cdot}01$, $V.R. = 21{\cdot}32$, $d.f. = 3, 6$.) Tapping to target words in the secondary, unattended message, however, was considerably more disruptive, causing over 30 per cent. errors and omissions in repeating the words coincident with or succeeding them in the primary message. (Here the variance ratio was significant, $p < 0{\cdot}001$, $V.R. = 29{\cdot}90$,

d.f. = 3, 6.) In an overall analysis of variance including both primary and secondary message and excluding "other" words (which were common to both) the difference in interference with the primary and secondary message was significant ($p < 0{\cdot}001$ *V.R.* = 156·9, *d.f.* = 1, 4). When subjects missed tapping, the results were reversed. In the primary message 76·5 per cent. target words which received no tapping response also failed to be repeated, as did about 20 per cent. of the neighbouring words. Two explanations are possible: either these words were simply not perceived, or the two responses competed so strongly that neither could be made overtly. Perhaps the most interesting point is that missed target words in the secondary message caused no increase whatever in interference with the repeating response, strongly suggesting that they were not identified at all. This is consistent with a perceptual limit to subjects' capacity which was fully occupied by receiving the primary message.

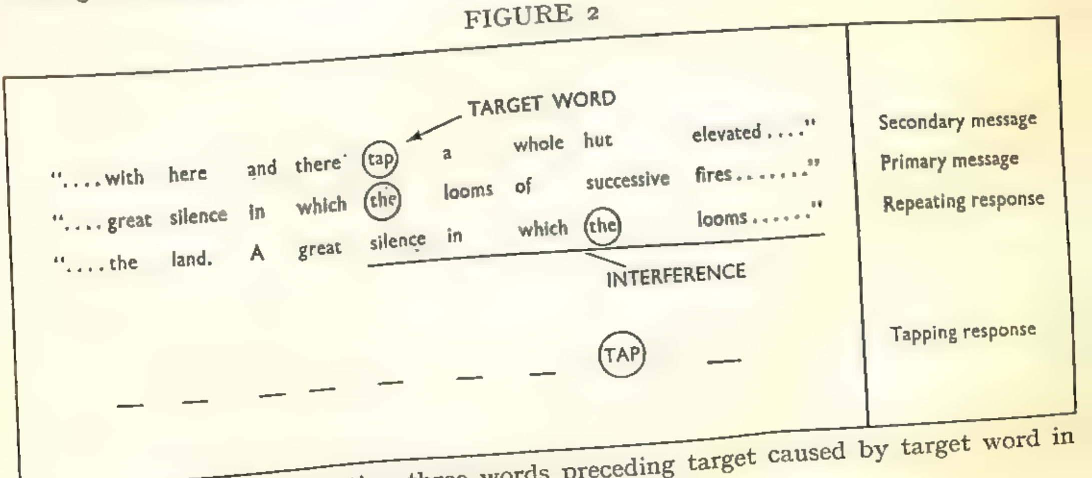
TABLE V

INTERFERENCE WITH REPEATING RESPONSE CAUSED BY TARGET WORDS

Class of word	*Per cent. errors and omissions when tapped*			*Other words*	*Per cent. errors and omissions when missed tapping*		
	3 Before	*Target*	*5 After*		*3 Before*	*Target*	*5 After*
Primary message	6·2	11·1	10·8	7·4	20·0	76·5	19·1
Secondary message	19·2	36·8	31·8	7·4	7·0	8·3	7·0

A possible difficulty with this interpretation is the high proportion of errors and omissions in repeating the three words before the secondary target, when this received a tapping response (a mean of 19·2 per cent. per word). Two explanations are possible: (1) it might be due to shifting of attention, so that the secondary message was actually functioning as primary message on those occasions where the target word was heard. (2) Since the mean ear-voice lag in repeating is three words, the subjects' attention might have remained on the primary message until the target word occurred, but its occurrence somehow erased the three previous primary message words or prevented the repeating response, perhaps by causing a subsequent

FIGURE 2



Interference with repeating three words preceding target caused by target word in secondary message.

shift of attention (see Fig. 2). To decide whether subjects had switched their attention *before* the target word occurred, we counted the omissions and errors in the preceding words, six to three before the target. The interference here was 8·2 per cent. which does not differ significantly from the 7·4 per cent. of the "other words." (This contrasts with 59·3 per cent. errors and omissions on each of the same words on those few occasions where the subject reported shifting his attention before hearing the target word.) It seems then that the 8·1 per cent. of secondary target words which received a tapping response were perceived despite the subjects' lack of attention, but that once these words had "got through" they interfered with repeating responses and perception of the following few words of the primary message.

Latencies. The response latencies for tapping and repeating were measured for the first nine subjects in Group A who received each of the three tape-recordings. We counted the number of primary message words intervening between the target word and the response, which had been recorded simultaneously on the second track of the tape-recorded results. For comparison the latency of the repeating response to a number of non-target words was also measured; in each case the word selected occurred five before the target word. The mean latencies are given in Table VI.

TABLE VI

RESPONSE LAGS, MEASURED IN WORDS

Response	*Tapping*	*Repeating target word*	*Repeating word 5 before target*
Primary message mean . .	2·9	3·1	3·2
Secondary message mean . .	3·0	—	—

To get some further indication whether the two responses were being organized simultaneously or successively, the latencies of each individual pair of responses to the same target word were subjected to t tests and to product-moment correlations. The t tests showed significant differences for only 11 out of 27 subjects; six showed shorter tapping than repeating latencies and five showed the opposite. The mean correlation over the different subjects was $r = 0·80$ ($d.f. = 1100$, $p < 0·001$).

Masking experiment. The mean percentage of target words detected by Group B in the masking experiment was 55 per cent. for the primary message and 52 per cent for the secondary message. These do not differ significantly, so there appears to have been no bias in the actual tape-recording favouring the primary message. Does external noise have an effect similar to that of inattention? The numbers of taps to each target word in the attention experiment were correlated with those in the masking experiment. Correlating the taps for each occurrence of each target word in the two conditions, the overall correlation was $r = 0·39$ for the primary message ($d.f. = 154$, $p < 0·001$) and $r = 0·18$ for the secondary message ($p < 0·05$). However, part of this correlation might have been due to the independent variables of verbal context and particular target words, both of which might affect the attention and the masking scores in the same way. When the correlations were calculated separately for each type of target word in and out of context and then averaged, the mean r for the primary message was still significant ($r = 0·27$, $d.f. = 138$, $p < 0·01$), but for the secondary, unattended message it was completely insignificant ($r = -0·06$). The intensity, clarity or whatever other factors increase intelligibility under masking were also somewhat helpful in increasing tapping responses to target words in the attended message, but seem to have had no effect in facilitating detection

of unattended words. The same conclusion emerged from the objective measures of intensity: the difference in intensity between target words and other words did not correlate at all with the number of tapping responses they evoked (the values of r for the different target words in the first recording were 0·00, 0·26, 0·15, −0·08 and 0·46 for the primary message, and 0·12, −0·06, −0·32, −0·06, −0·20 for the secondary message). None of these correlations is significant. This was rather a surprising result, which will be discussed later.

Cerebral dominance. Finally we come to the results for the control Group D, relating to cerebral dominance and left-right asymmetry. Table VII gives the

TABLE VII

EFFECT OF LEFT-RIGHT DIFFERENCES ON TAPPING RESPONSES

		Parts of face		*Point*		*Tired*		
	Tap	*In context*	*Out*	*In*	*Out*	*In*	*Out*	*Mean*
Primary message	Left 90·0	90·0	80·0	93·3	90·0	93·3	86·7	89·0
	Right 93·3	90·0	76·7	86·7	96·7	96·7	96·7	91·0
Secondary message	Left 10·0	10·0	6·7	6·7	3·3	33·3	16·7	12·4
	Right 0·0	10·0	10·0	3·3	0·0	6·7	6·7	5·2

mean percentages of correct tapping responses in the different conditions for the left and right ears. On analysis of variance, the difference between left and right ear target words was completely insignificant for the primary message, but target words on the left ear in the secondary, unattended message received significantly more taps than those on the right ear ($p < 0·025$, $d.f. = 1, 18$, $V.R. = 7·67$). The difference is mainly due to the passage containing "Tired," in context, and the interaction between target words and left-right difference was significant ($p < 0·025$, $d.f. = 2, 18$, $V.R. = 6·23$). This shows that the choice of left ear for the unattended message in the main experiment was, if anything, favouring responses to the secondary target words, probably because the primary task of repeating was easier for messages on the right ear and so left more spare capacity for the secondary task.

The efficiency of the repeating response also showed differences between left and right ear messages. The percentage of errors and omissions in repeating the "other words" (i.e. not target, three before or five after) was 10·0 per cent. for the left ear and 6 per cent. for the right. This difference was significant on analysis of variance ($p < 0·01$, $V.R. = 19·0$, $d.f. = 1, 9$). There was no significant increase in left-right asymmetry when subjects repeated the target words, the three before and the five after.

Discussion

Perceptual or response limit?

Our first aim was to discover how far attention is limited by a restriction on perceptual capacity and how far it is limited by the number of responses which can be simultaneously organized. The difference in tapping responses to primary and secondary messages gives the answer, which overwhelmingly favours a perceptual

limit with a filter selecting before the two messages are fully analysed, as in Figure 1*a* rather than 1*b*. Since both stimulus and response are identical for the primary and secondary message, it is difficult to argue for a difference in importance, in response load or in response bias, and the result seems best explained on the assumption that the secondary target words are much less likely to be identified than the primary ones. There is also some degree of response competition shown by the number of occasions when one response to primary target words was given but not the other (about 19 per cent.). But in this task the response competition is much less dramatic than the perceptual competition.

Lawson (1966) has recently repeated this experiment with an interesting difference in the task and very different results. Instead of words as her signals to tap, she used brief tones or pips. The difference between responses to the primary and secondary messages almost completely disappeared. It seems that analysis of simple physical signals precedes both the selective filter and the analysis of verbal content in the perceptual sequence, that the bottleneck in attention arises chiefly in speech recognition, where of course the information load is usually much higher. To confirm the belief that the verbal content of the secondary message in the present experiment was not being analysed, we find no evidence whatever of interference from secondary target words when these received no tapping response.

On the other hand, in those cases where the subject did tap to the target word in the secondary message, this caused significantly more interference with the repeating response than tapping to the primary message, (more than 30 per cent. errors or omissions on each of the target and five succeeding words compared with 11 per cent. for primary target words). In both cases the same two responses are made; the only difference is that there are two words to identify when the target is in the secondary message and only one when it is in the primary message. Thus the difference in interference again supports the hypothesis of a perceptual limit. Two further points support the idea that most of the secondary message was not perceived: (1) the fact that the verbal context of the target words probably does not facilitate responses as it does in the primary message; (2) subjects' failure to distinguish the homophones of "Right" in the secondary message. The only verbal factor we expected would have an effect on taps to the secondary message was the information content of the classes of words compared to the single words, since these would impose a greater load on the limited perceptual capacity. This did affect performance: the classes of words received significantly fewer tapping responses than the single words with one main meaning.

Stimulus variables in primary message. In contrast, many of the stimulus variables affected perception of the primary message (as shown by the number of target words receiving at least one of the two responses). The verbal context had a marked effect increasing both tapping and repeating responses. Increased size of the target word ensemble led to significantly poorer performance, decreasing the per cent. correct and increasing the latency, particularly when the target words were out of context. When they were in context, the difficulty due to increased ensemble was almost cancelled out by the high transition probabilities. These findings confirm once again that our perceptual capacity is limited at least partly by the information content of stimuli presented. In this experiment, however, the information limit was shown only when the class of target words included different phonetic patterns and not when its members varied only in meaning. The two stimulus variables which also affected responses directly were the predictable or random insertion of target words in the context of the passages and the grammatical difference between lexical and function words. The proportion of target words receiving one response

but not the other was much higher for target words out of context (24·5 per cent. compared with 8·2 per cent.). Here we have a dissociation between the two ways of increasing information—increasing the size of the ensemble and decreasing the transition probability. The former appears to affect only the word's chance of being correctly perceived, while the latter affects also the selection of responses. For the function words in context, there were abnormally few tapping responses compared with other grammatical classes of target words. The relevance of these differences is discussed below.

Organization of two competing responses

The third problem raised in the Introduction is the relation between two competing responses made to the same perceived signal. At what stage in the sequence of perceptual decisions are the responses selected, and does this affect the degree to which they interfere with one another? The present results give no conclusive answers to these questions, but they may give some indications.

Firstly we compared the latency of the tapping and the repeating responses to the primary target words. The mean latency of the tapping response was shorter, by about 80 millisec., but t tests on the pairs of latencies to individual target words showed significant differences for only 11 out of 27 subjects. The correlation between the two latencies was high for nearly all subjects, which makes it unlikely that the word was being analysed by two independent systems for the two separate responses. The result gives little evidence either for or against serial programming of the tapping and repeating responses.

Secondly we compared the characteristics of the primary target words which affected the repeating and the tapping responses, to get some indication of the stage in the perceptual sequence at which each was initiated. Table II shows that they followed a similar pattern on the whole. Both responses were worse with classes of words than single words, but not with words of many meanings or homophones. This could be taken as evidence that both responses are initiated together before the meaning is analysed. However an alternative explanation is possible: since the different meanings of "Fit," "Point," "Right" and its homophones are only brought out when the words are in context, any resultant decrement due to the increased ensemble of meanings might be cancelled out by facilitation from the verbal context. Even the digits, colours, etc., were only about 5 per cent. worse than the single words, when they were in context. Further evidence about the tapping response comes from the condition where tapping to homophones was incorrect: when subjects were asked to tap only to "Right" they tapped mistakenly to 30 per cent. of the homophones in the primary message, compared with 94 per cent. when the homophones were correct. In this case, subjects were certainly identifying the meaning of two-thirds of the target words and tapping to the speech sound for one-third. However this condition differed from the others in that the instructions to avoid tapping to homophones stressed the importance of analysing meaning before tapping. Thus the 30 per cent. mistaken taps may be more significant than the 64 per cent. of homophones avoided.

The variable which causes the most striking divergence between the tapping and the repeating responses is the verbal context. While the lack of context severely impaired the repeating response, reducing the correct words by 29 per cent., the decrement for the tapping response was only 8 per cent. When words were out of context, tapping was consistently better than repeating. This suggests that the tapping response was not dependent on the same high level of analysis as the repeating response. While subjects relied on the verbal context to give the meaning

of target words before they could repeat them back, this was not essential for the tapping response. On the other hand the tapping response was facilitated a little by the verbal context; the improvement of 8 per cent. was statistically significant. Two explanations are possible: either the tapping response was made, in some cases at least, to the verbal unit rather than the speech sound, or the effect was entirely due to the greater ease of repeating words in context, which left more spare response capacity for tapping.

The only other points in Table II at which the two responses diverge are the target words "From" and "But," which in context are much worse at evoking a tapping than a repeating response, and the target word "Tap" where the tapping response is much better, presumably because of the high stimulus-response compatibility. On the assumption that the tapping response is triggered by the speech sound, we might explain the difficulty of "From" and "But" in context by saying that the speech sounds are less distinct: with sentence intonation, these function words will seldom carry as much stress as the lexical items. If we assume full verbal analysis, the explanation may be that "From" and "But" are not perceived as distinct functional units in the same way as the lexical words, but simply as part of the syntactical unit or phase in which they occur. The repeating response could mirror the whole phrase, while the tapping response required the subject to isolate the particular words "From" or "But" within the phrase. In this case the verbal context actually makes the tapping response more difficult.

In conclusion then, we have no convincing proof either that the responses are always organized successively or that they are always initiated at the same point in the perceptual hierarchy from speech sound to meaning. However, any tests which were not entirely ambivalent favoured the alternative of serial programming at different levels of analysis: the slightly shorter latency for tapping, the much greater difficulty of repeating than of tapping to words out of context and the failure to identify a third of the homophones before tapping, all suggest that the tapping response was triggered at a lower level than the repeating response. Any evidence which might suggest simultaneous organization can also be explained in a way consistent with serial organization. If the serial hypothesis is correct, it might account for the relative lack of response competition found in this experiment.

Stimulus-response compatibility

The third question raised in the Introduction was how far compatibility of stimuli and responses might allow them to bypass the limited capacity of selective attention. The result differed for the primary and the secondary messages: while the target word "Tap" in the primary message received significantly more and quicker tapping responses than the corresponding adjectives "Tall," "Hot" and "Tired" out of context, in the secondary message there was no difference at all. This is quite consistent with the model of selective attention controlled by a perceptual filter. Stimulus-response compatibility can be interpreted as the high conditional probability of a particular response given a particular stimulus—a reduction in response but not stimulus uncertainty. One would therefore expect it to affect subjects only once the target word had been identified: it might then be expected in the primary, attended message to facilitate the tapping response rather than the repeating response, which it does, and perhaps also to decrease the response competition, which it does not (tapping to the target word "Tap" caused if anything more interference with the repeating response than the other target words). If we are right that little of the secondary, unattended message is being perceived, the target word "Tap" should gain little benefit from its compatible response.

Nature of perceptual filter

The last point to discuss is the nature of the perceptual "filter" which so drastically reduces recognition of the secondary message. The subjects did hear a few of the target words from the secondary message, but showed no evidence of hearing anything more. Treisman (1960, 1964) suggested a modification of Broadbent's original model, based on an analogy with the signal detection theory of sensory thresholds (Tanner and Swets, 1954). If the filter reduced the signal-to-noise ratio of unattended messages rather than blocking them completely, words which were highly important or relevant to the subject might still be perceived despite this attenuation, provided that the criteria for detecting them were sufficiently low. This would have the biological advantage that unattended messages could be monitored for any important signals, without at the same time much increasing the load on the limited capacity available for speech recognition. Broadbent and Gregory (1963) measured changes in signal strength and criterion when the subject was attending to and away from a tone masked by noise and obtained results consistent with this suggestion. It explains how in the present experiment subjects were able to hear about 8 per cent. of the unattended target words, when told in advance what these would be.

However a point which seems at first sight unexpected is the lack of correlation between the particular words which were detected in the attention and in the masking experiments. If the filter has the effect of reducing the signal-to-noise ratio of unattended messages, one might expect some parallel with the effects of an external masking noise, in that those features of target words which made them likely to survive the one would also make them likely to survive the other. However there is an important difference between these conditions which might explain the lack of correlation: in the attention experiment subjects are occupied with the primary message, and the degree of attenuation of the secondary message affected by the filter may vary with the load on attention imposed by the primary message. This will probably fluctuate from moment to moment, with the predictability or difficulty of the words, the rate at which they are spoken and so on. If so the signal-to-noise ratio of unattended words will fluctuate randomly in relation to particular target words and this random variation may swamp any correlation with the effects of the constant external noise used in the masking experiment.

Another prediction from this model of the attention process is that there should be some false positives, related in sound to the target words in the unattended message. There were a few of these: subjects tapped to a total of 11 non-target words in the whole experiment (compared to 232 target words detected), and all of them were similar sounds such as "both" for "boat," "light" for "night," "at" for "but" (except one which was "face" for "any part of the face," a semantic error). It is not possible to work out an exact false positive rate, since one does not know the number of words at risk. As a very rough guide, excluding the passages with digits, colours or parts of the face, there were an average of 7·6 words per passage which shared two phonemes with the target word, as did all the actual false positives made. This would give a false positive rate for all passages with single target words, of 0·17 per cent. compared to a hit rate of 11·2 per cent. In the masking experiment the corresponding rates are 3·4 per cent. false positives and 58·3 per cent. hits, assuming that the same mean of 7·6 words per passage are at risk. Finally for the primary message in the attention experiment, the rates are 0·19 per cent. false positives and 89·7 hits.

It is tempting to look up the corresponding values given by signal detection theory for d′, the signal strength, and β, the criterion, in each of these tasks. However the assumptions underlying this use of signal detection measures are questionable: (1) we

assume a central continuous dimension of evidence determining perception of particular target words, along which the actual target and "noise" words vary in similarity; (2) we assume that their distributions on the similarity dimension are normally distributed, overlapping and with approximately equal variance; (3) we take a rather arbitrary number of 7·6 for the set of "noise" words in any passage; (4) finally we have pooled the results for all subjects and all target words in the hit and false positive rates given above. (It seems reasonable to exclude the digits, colours and parts of the face, since the criteria may well differ where a class of different sounds must be detected and where a single sound is the target.) These assumptions mean that little weight can be given to the absolute values for d′ and β, but since the assumptions are constant for all three conditions, it may be worth looking at the directions in which the values change as we change the task variables. On the general model we suggested, for the masking condition we should expect a low d′ (because of the added noise) and a relatively low criterion (since this is the subject's only task); for the secondary message in the attention task, we predict a low d′ (due to the reduction in signal-to-noise ratio introduced by the filter) and a relatively high criterion (because tapping to target words was a secondary task competing with the primary repeating response); finally for the primary message in the attention task we expect a high d′ (since the message is unmasked and receiving full attention) and again a high criterion

TABLE VIII

Changes in Signal Strength and Decision Criterion in the Different Experimental Conditions

	Masking (*primary message masked by noise, primary response*)	*Inattention* (*secondary message, secondary response*)	*Attention* (*primary message, secondary response*)
d′	2·0	1·8	4·2
β	5·1	34	30

(because tapping is the secondary response). Table VIII gives the values of d′ and β obtained. They are quite consistent with the predictions.

We should like to thank the Medical Research Council for supporting this research, Mr. C. Turner who helped carry out a pilot experiment, and Dr. M. Treisman for his helpful criticisms.

References

Broadbent, D. E. (1958). *Perception and Communication.* London: Pergamon.

Broadbent, D. E., and Gregory, M (1962). Donders' b- and c- reactions and S–R compatibility. *J. exp. Psychol.*, **63**, 575–8.

Broadbent, D. E., and Gregory, M. (1963). Division of attention and the decision theory of signal detection. *Proc. Roy. Soc. B.*, **158**, 222–31.

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *J. accoust. Soc. Amer.*, **25**, 975–9.

Davis, R., Moray, N. P., and Treisman, A. M. (1961). Imitative responses and the rate of gain of information. *Quart. J. exp. Psychol.*, **13**, 78–89.

Deutsch, J. A., and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.*, **70**, 80–90.

Lawson, E. A. (1966). Decisions concerning the rejected channel. *Quart. J. exp. Psychol.*, **18**, 260–5.

Leonard, J. A. (1961). Choice reaction time experiments and information theory. In Cherry, E. C. (Ed.) *Information Theory*. London: Butterworths, p. 137–46.

Moray, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quart. J. exp. Psychol.*, **11**, 56–60.

Moray, N., and Jordan, A. (1966). Practice and compatibility in two-channel short term memory. *Psychon. Sci.*, **4**, 427–8.

MORAY, N., and TAYLOR, A. (1958). The effect of redundancy in shadowing one of two dichotic messages, *Language & Speech*, **1**, 102–9.

MOWBRAY, G. H. (1960). Choice reaction times for skilled responses. *Quart. J. exp. Psychol.*, **12**, 193–9.

MOWBRAY, G. H. (1964). Perception and retention of verbal information presented during auditory shadowing. *J. acoust. Soc. Amer.*, **36**, 1459–64.

REYNOLDS, D. (1964). Effects of double stimulation: Temporary inhibition of response. *Psychol. Bull.*, **62**, 333–47.

TANNER, W. P., Jr., and SWETS, J. A. (1954). A decision-making theory of visual detection. *Psychol. Rev.*, **61**, 401–9.

TREISMAN, A. M. (1960). Contextual cues in selective listening. *Quart. J. exp. Psychol.*, **12**, 242–8.

TREISMAN, A. M. (1964). Selective attention in man. *Brit. Med. Bull.*, **20**, 12–6.

Manuscript received 28th July, 1966.

OBSERVING RESPONSES AND UNCERTAINTY REDUCTION

BY

J. STEINER*

From Psychological Laboratory, Cambridge

Rhesus monkeys and baboons were placed in an observing response situation where on any trial they could work for food in the presence of an imposed stimulus or make an observing response by pressing a lever to present an alternative stimulus. The reinforcing properties of the alternative stimulus were assessed by placing the observing response on a progressive ratio schedule and were found to vary as a function of the difficulty of the imposed discrimination and the degree to which the imposed discrimination had been learned. The results are consistent with the hypothesis that the reinforcing strength of the alternative stimuli varies with the amount of uncertainty they reduce. Eliminating the response requirement showed that the effect was not simply due to the fact that informative stimuli permitted a saving of responses.

INTRODUCTION

The observing response paradigm introduced by Wyckoff (1952) is a useful one with which to study the reinforcing properties of discriminative stimuli. His pigeons pecked a translucent key for grain on a schedule made up of 30 sec. trials, half of which were positive, terminating in reinforcement, and half negative, terminating without reinforcement. The pigeons could choose either to peck with the key illuminated by an imposed stimulus which was white on both positive and negative trials, or to make an observing response by standing on a pedal which resulted in the presentation of one of a pair of alternative stimuli. If the trial happened to be a positive one the key became red, while if it was negative it became green. Wyckoff found that observing responses increased as the birds began to discriminate between the alternative stimuli, i.e. as they began to peck more on red than on green trials. If the colours were arranged to no longer correlate with the type of trial, observing responses decreased, while if the stimuli were reversed, observing decreased and then increased again as the birds began to respond appropriately to the reversed stimuli.

The observing response in no way altered the probability of food reinforcement, but it is not true to say as Wyckoff does that the stimuli provided "only information," implying that no material advantage accrued. Although they could not obtain more reinforcement by making observing responses, the pigeons could and did distribute their responses more economically with the aid of the alternative stimuli. Because of the fact that they paused, or responded at a lower rate in the presence of the negative stimulus, the overall response cost per reinforcement was lower in the alternative than in the imposed condition. The stimuli, therefore, resulted in some material gain which could account for their reinforcing properties.

Indeed, in most situations, the subject makes use of the information provided by stimuli. Sometimes the information enables reinforcement to be obtained which would otherwise be unobtainable, sometimes responses are saved, and sometimes it enables the frequency of reinforcement to be increased. The latter was the case in the experiments on observing responses by Kelleher (1958), Kelleher, Riddle

* Now at Department of Psychiatry, Maudsley Hospital, London.

and Cook (1962) and Zeigler and Wyckoff (1961) where the frequency of reinforcement was higher in the alternative than in the imposed condition, and in these situations it is not difficult to see the use to which the information is put.

This is less clear, however, in an experiment by Prokasy (1956) which differed from the others in that no specific food producing response was required. Prokasy's rats were given a choice between the two arms of a T-maze. At the end of each arm they entered a delay box where they waited for 30 sec. before entering a goal box in which food was found randomly on 50 per cent. of trials. The delay boxes were either black or white and the only difference between the two arms of the maze was that on one side the colour of the delay boxes was correlated with reinforcement and non-reinforcement respectively, while on the other side the colours and reinforcement were randomly related. The animals were given equal experience of both arms by a procedure of free and forced trials, and as training progressed they showed a preference for the correlated side.

A response to the correlated side can be considered analogous to an observing response since it resulted in exposure to discriminative stimuli which signalled reinforcement and non-reinforcement. The stimuli on the non-correlated side provided no information of the type of trial and hence are analogous to the imposed stimuli of Wyckoff's experiment. In Prokasy's experiment, however, the two choices were equated for both frequency of reinforcement and response cost per reinforcement, and it is not clear what the animal gained by seeking the discriminative stimuli.

Somewhat analogous behaviour is commonly seen when people are waiting for a bus or train. Here observing responses are frequently made to see if the bus is coming or not, although this does not increase the probability of its arrival. Sometimes, it is true, that when the bus is coming preparatory responses might ensure a better seat, or a taxi might be taken if the bus is not in view, but for the most part the advance information as to whether the bus is coming or not, has no obvious usefulness. Similarly we are eager to know our examination results at the earliest possible date; we stay up to the early hours of the morning to hear election results, and few of us would resist the temptation to look at our presents if left alone in the house on the day before our birthday (Berlyne, 1960). The information in these situations usually results in no material gain, but does involve a reduction of uncertainty. Discussing this topic Berlyne comes to the conclusion that "Uncertainty is manifestly one of the burdens that the human frame is least equipped to stand."

When the usefulness of the information is obvious it is not difficult to see why informative stimuli should be reinforcing but the above examples suggest that other sources of reinforcement strength should be looked for. It is possible, for instance, as Berlyne (1960) suggests, that uncertainty is itself aversive because it leads to conflict and that informative stimuli are reinforcing by virtue of the fact that they reduce uncertainty.

Whatever the cause, there is plenty to suggest that a close relationship exists between the reinforcing properties of discriminative stimuli and the information they provide. The aim of the present experiments was first to explore this relationship further by introducing quantitative concepts, namely to see how the reinforcing *strength* of discriminative stimuli varied with the *amount* of information they provide or what is equivalent, with the amount of uncertainty they reduce: and second, to see if the reinforcing properties of the informative stimuli depended on the saving of responses which the information permitted. (A stimulus provides information only in so far as it reduces uncertainty, and the amount of information provided is determined by the amount of uncertainty reduced (Attneave, 1959; Garner, 1962).)

Experiment I

In all of the work on observing responses to date, discriminative stimuli have been present in the alternative condition and absent in the imposed condition. This need not be the case, however, and in this experiment an observing response was arranged to switch from one set of discriminative stimuli to another.

The set up was similar to that used by Wyckoff except that his single neutral imposed stimulus was replaced by one of a pair of stimuli which marked positive and negative trials respectively. The amount of uncertainty associated with the imposed stimulus could then be varied by increasing the difficulty of discriminating between the positive and negative imposed stimulus. As in Wyckoff's experiment the animals could either remain in the presence of the imposed stimulus or make an observing response which replaced this with a red stimulus on positive trials and a green one on negative trials, these alternative stimuli remaining easy to discriminate throughout the experiment.

Difficult discriminations have been considered to generate conflict in another context, namely in the work done on "experimental neurosis" where disturbed behaviour has been reported as stimuli were brought together progressively along a continuum (Pavlov, 1927; Broadhurst, 1960). The observing response in the present experiment can be viewed as a means of escape from such a discrimination, and the animals were carefully observed for signs of "neurotic" behaviour.

The difficulty of the imposed discrimination was manipulated by varying the difference in brightness between two stimuli and this formed the independent variable of the experiment. For the chief dependent variable, a measure of the reinforcing strength of the alternative stimuli was required and since previous techniques were considered unsatisfactory for a variety of reasons (Steiner, 1964) a new method making use of a progressive ratio schedule (Hodos, 1961) was developed. On each trial the animals could either remain in the presence of the imposed stimulus or press a lever which resulted in a switch to one of the alternative stimuli. The number of bar presses required, however, increased by a step of two following each successful switch. On the first occasion therefore, one bar press was sufficient, on the next three were required, on the third, five, and so on. Switching to the alternative stimuli thus required a greater cost, in terms of bar presses, as the number of switches increased. At the beginning of each session the number of bar presses required returned to one, and on each day the level was "titrated." The number of switches made was taken as a measure of the reinforcing strength of the alternative stimuli on that particular day.

Method

Subjects. The subjects were two Rhesus monkeys, M1 and M2, and one Baboon B1, all of which had been used in previous experiments in the same apparatus (Steiner, 1964). The animals were moderately food deprived but not reduced to any fixed weight. They earned most of their daily intake of food in the testing session where each reinforcement consisted of two pellets of chow, and were given a supplement in their home cages to make up a total of 100 gm. of chow daily as well as some fruit and vitamins. M2 required a greater amount of food to maintain his weight and received three pellets per reinforcement, making a daily total of 120 gm. of chow.

Apparatus. The animals faced a 3-in. square, opal glass, swinging panel which was hinged at the top and lightly sprung. Pushing this through approximately 15 degrees closed a microswitch and delivered food on the schedule described below. The imposed and alternative stimuli were back-projected on to the panel by slide projectors and coloured lamps.

The observing response lever attached to a post office key protruded into the box in a central position about 2 in. below the panel. Food reinforcements were delivered by a Gerbrands Universal Belt Feeder into a food tray situated about 3 in. above the panel.

The entire experiment was programmed automatically by a rack of relays, uniselectors, and clocks situated in the corridor outside the room. A Gerbrand cumulative recorder monitored panel pressing and indicated when a switch to the alternative stimulus occurred. The number of panel presses in the presence of each stimulus and the number of lever presses were also recorded on banks of counters.

Schedule of food reinforcement. The daily session consisted of 20 positive and 20 negative trials in a near random sequence, the only restriction being that no more than four trials of the same type could occur in succession. On positive trials panel pressing was reinforced by pellets of chow on a fixed interval 1 min. schedule with a limited hold of 5 sec. This means that the first response after 1 min. had elapsed was reinforced, but if no response occurred after 65 sec. had elapsed, the trial terminated without reinforcement. Responses were not reinforced on negative trials which terminated at the end of 65 sec. Without the limited hold the animals could have solved the problem by waiting until the minute was over, when the trial would terminate if it was negative but not if it was positive. On this schedule some responding continued on negative trials, perhaps maintained by superstitious association with the termination of the trial. Negative trials were, therefore, arranged not to terminate within 1 sec. of a panel press and were thus occasionally slightly prolonged.

Stimuli. The imposed stimulus was provided by one projector on positive trials and the other on negative trials. The brightness of each was adjusted by passing the light through crossed polaroids attached to the front of each projector and these were calibrated by means of an SEI exposure photometer. The difficulty of the discrimination was varied through five stages which are shown in Table I.

TABLE I

BRIGHTNESS OF STIMULI IN LOG FT. LAMBERTS

	Stage				
	A	*B*	*C*	*D*	*E*
Positive stimulus	2·5	1·5	1·4	1·3	1·3*
Negative stimulus	1·0	1·1	1·2	1·3	1·3*

* Both projectors on.

In stage D the brightness of both stimuli was measured as 1·3 log ft. lamberts but the positive one did appear very slightly brighter. In stage E, therefore, the stimuli were made indiscriminable by having both projectors on in all trials with the polaroids adjusted so that together they produced a brightness of 1·3 log ft. lamberts.

If the animal made a successful switch to the alternative stimuli the projectors were turned off and the panel was illuminated by a red light if the trial was positive and a green one if it was negative.

As an index of the performance on the discrimination a discrimination ratio (D.R.) was calculated from the mean number of responses per trial in the presence of the positive and negative stimuli in the following way:

$$\text{D.R.} = \frac{\text{mean responses in S}^+}{\text{mean responses in S}^+ + \text{mean responses in S}^-}$$

The D.R. was expressed as a percentage and varied from 100 per cent. if all responses were made to the positive stimulus, to 0 per cent. if they were all to the negative stimulus. When response rates were equal in the two stimuli, i.e. when discrimination was at a chance level, the D.R. should have been equal to 50 per cent. but because of the slighty greater length of negative trials, it was mostly a little less than this.

The D.R. for the alternative discrimination remained close to 100 per cent. throughout the experiment. In the imposed discrimination the D.R. varied with the difficulty of the discrimination and was calculated for each session.

Observing response schedule. On each trial the animals could either remain in the presence of the imposed stimulus or switch to the alternative stimulus, but they could

not alter the type of trial. On a negative trial the dark stimulus was present and switching would alter this to the green, while on a positive trial the bright stimulus could similarly be changed to the red.

The number of bar presses required to switch from the imposed to the alternative stimuli (referred to as the ratio requirement) increased by a step of two following each successful switch. If the animal failed to meet the required ratio on the bar, it remained in the presence of the imposed stimuli and the requirement on the bar did not increase. The observing responses could be made at any point in the trial, and once produced, the alternative stimuli remained on for the rest of the trial.

In order to prevent superstitious association between observing responses and food reinforcement, the trial did not terminate within 5 sec. of a bar press. Bar pressing did sometimes prolong the trial, particularly early in training, but later the animals usually made the required number of bar presses quickly at the beginning of the trial if they pressed at all.

The number of switches made each day was taken as a measure of the reinforcing strength of the alternative stimuli. Because of the progressive ratio schedule, this was directly proportional to the number of lever presses the animal was prepared to make to switch.

Procedure. The experiment was begun with stage E of the imposed discrimination which was followed by stage A in which the discrimination was easy and then through stages B, C, D, and finally stage E once more. The animals remained at each stage until it seemed from inspection of the record that no further change was likely, and this took from 6 to 17 days.

Results

Although all of the animals showed peculiarities in their behaviour, nothing suggestive of "experimental neurosis" was observed throughout the experiment.

The animals switched to the alternative stimuli for a number of trials and then, when the ratio requirement on the bar became too high, remained in the presence of the imposed stimuli. Frequently, however, they resumed bar pressing after having failed to switch for some trials. Bar pressing usually began as soon as the trial commenced, but sometimes the animals would panel press for a while, and then change their minds and begin bar pressing. If the green or an easily discriminable dark negative stimulus was present they usually turned away from the panel and engaged in alternative behaviour, but occasionally, particularly after a run of non-reinforced trials they would make some panel presses to the negative stimulus.

The daily performance for each animal is plotted in Figure 1 which is explained in the legend.

(i) Stage E. In this stage the imposed stimuli were indiscriminable and the animals performed at a chance level. M1 and M2 switched on about 10 trials, thus meeting a ratio of 19 responses on the lever, and B1 switched on 15 to 20 trials which required up to 39 lever presses.

(ii) Stage A. Positive and negative trials were easily discriminable at this stage and performance reached a high level, very few panel presses being made on negative trials. The number of switches fell in all animals, although the change was small in the case of M1. In the case of M2 the fall was preceded by a transient rise. Bar pressing, however, was not completely eliminated, and the animals almost invariably switched on the first trial and often on the next few trials.

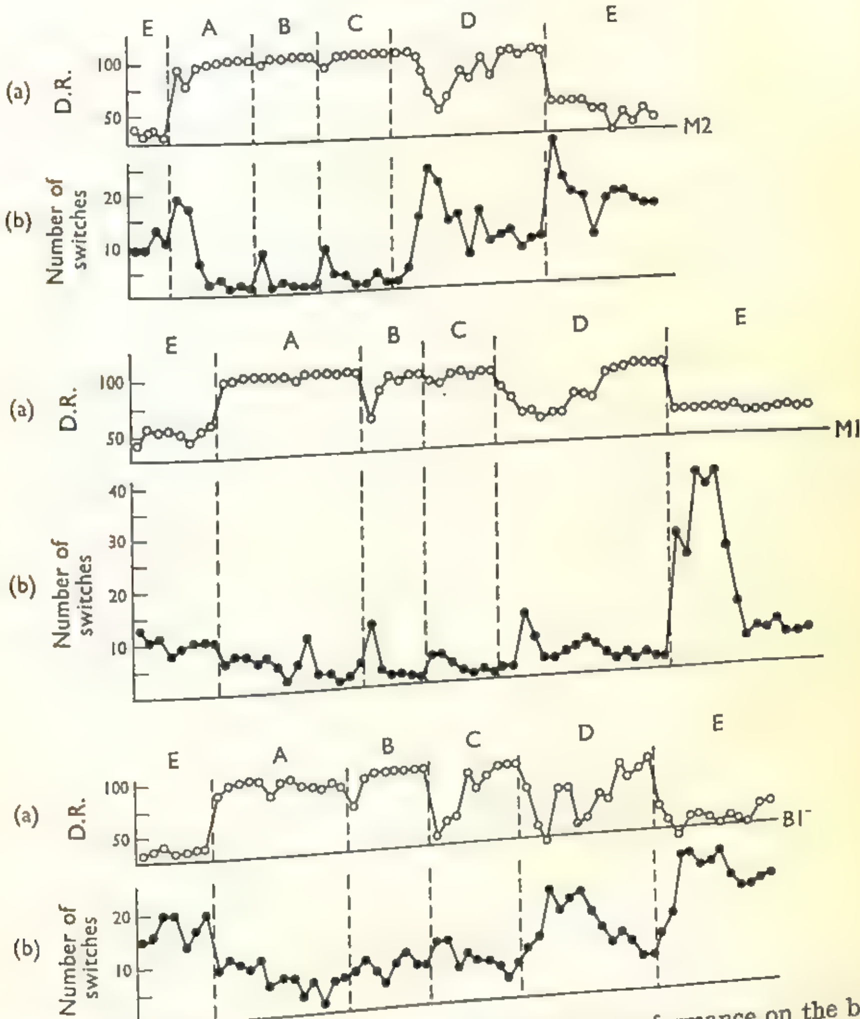
(iii) Stages B, C and D. At the beginning of each subsequent stage of difficulty, performance on the discrimination tended to fall, but in all cases subsequently recovered. The falls became increasingly severe and prolonged with each successive stage, and in all cases took the form of an increased number of panel presses in negative trials, the rate on positive trials remaining approximately the same.

With each drop in performance, there was an increase in the number of switches made to the alternative stimuli, which tended to return to its previous level as

performance improved. These changes became greater with each successive stage of difficulty.

(iv) Stage E reintroduced. When stage E was reintroduced, discrimination ratios fell once again to chance levels, and the number of switches increased. M1 reached extremely high levels on the first 6 days, meeting ratios as high as 75 responses on the bar, but on the first five of these days, he hardly pressed the panel at all in the imposed condition, where he behaved as he had done on negative trials in earlier stages, frequently shaking his head between his legs. On the sixth day, he

FIGURE 1



Daily results in Experiment 1. Section (*a*) shows the performance on the brightness discrimination as a discrimination ratio (D.R.). Section (*b*) shows the number of switches made to the alternative stimuli. The stages of difficulty are marked at the top of the figure.

began to panel press in the presence of the imposed stimuli and the number of switches fell to a level slightly lower than they had been on the first presentation of stage E. B1 also switched slightly more than he had done previously but eventually maintained a level similar to his previous one. At first M2, like M1, did not panel press in the presence of the imposed stimuli and met very high ratios on the bar. On the

third day, however, he resumed pressing and the number of switches fell to a level somewhat higher than it had been when this stage had been presented at the beginning of the experiment.

Discussion

The situation in stage E was similar to that used by Wyckoff and the results confirm his findings. The alternative stimuli were sufficiently reinforcing to maintain switching on a progressive ratio schedule even though an observing response led to no increase in the frequency of reinforcement.

In stage A the imposed stimuli were easily discriminable and the number of switches fell to a low level. As each new stage of difficulty was introduced, the performance on the discrimination tended to fall, and the number of switches to the alternative stimuli tended to increase. Both these changes were transient, however, and as the animals became familiar with the new, more difficult pair of stimuli, the performance improved and the preference for the alternative stimuli returned to its previous level. When the final indiscriminable stage was once more introduced, switching increased markedly and then fell once more to a level similar to that which had been maintained in the indiscriminable stage at the beginning of the experiment.

The alternative stimuli, therefore, were most reinforcing when we would expect uncertainty to be greatest, namely when no discrimination was possible between the imposed stimuli, and temporarily following each increase in the difficulty of the discrimination. The results are thus consistent with the hypothesis that the reinforcing strength of the alternative stimuli increased with the amount of uncertainty they reduced.

Experiment 2

In this experiment the change in reinforcing strength of the discriminative stimuli was followed while learning was taking place.

Procedure. The general arrangement was identical with that used in Experiment 1. Initially a white imposed stimulus was present and the reinforcing strength of the alternative red-green stimuli was assessed as before. When behaviour had stabilized, a blue-yellow discrimination replaced the white imposed stimulus and the reinforcing strength of the alternative stimuli was assessed as it was learned. Next a second discrimination between a cross and a circle replaced the blue-yellow stimuli until this too was learned, and finally the initial white stimulus was imposed once more.

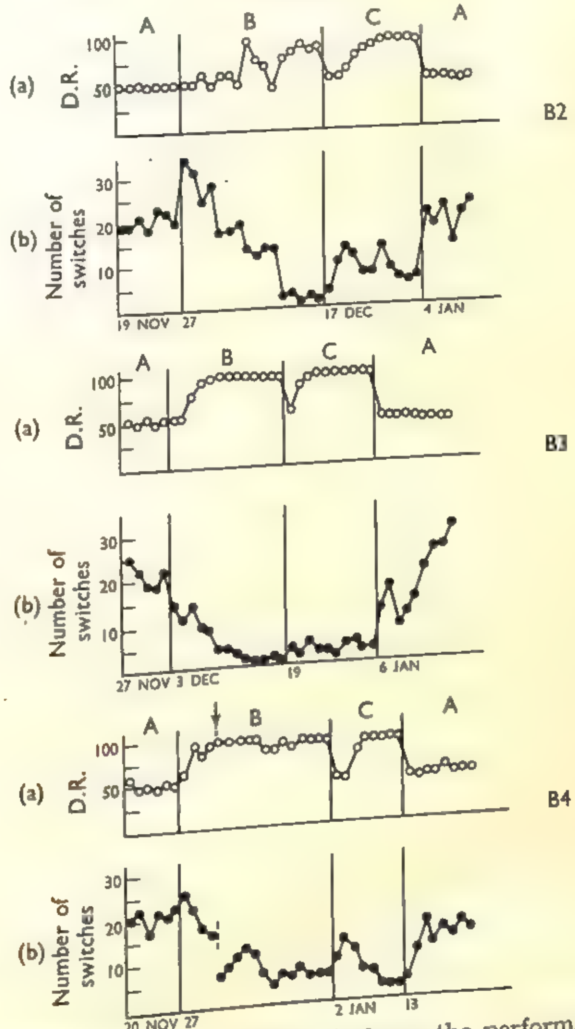
Subjects. Three new experimentally naive baboons B2, B3 and B4 were used. Pre-training was done in gradual stages and lasted about 5 weeks. With the white imposed stimulus present the step by which the ratio requirement on the bar increased after each successful switch was adjusted so that the animals switched on approximately 20 of the trials and remained in the imposed condition for the remaining 20. For B2 the step was three, for B3 and B4 ten. These steps were considerably higher than those required by the animals in Experiment 1 and the reason for this was not clear. B2 and B3 developed very rapid rates of bar pressing and sometimes reached ratios as high as 250 responses even though the end of the trial was delayed when this occurred since over a minute was required.

Stimuli. The blue, yellow, cross and circle stimuli were made up as slides and inserted in the projectors. For the white stimulus no slide was used and one projector illuminated the panel on all trials. As in the previous experiment discrimination ratios were calculated from the number of panel presses made in the presence of each stimulus. As before the negative stimulus periods were somewhat longer than the positive ones because they did not terminate until the limited hold was over. In this experiment, to make the rates in positive and negative trials comparable, only those panel presses made in the first 60 sec. were used in the calculation of the D.R. Chance performance, therefore, resulted in a D.R. of 50 per cent.

Results

The results are shown in Figure 2. In section (*a*) the D.R. in the imposed condition is plotted. Performance on the alternative red-green discrimination remained relatively good and is not shown. When the imposed stimulus was white no discrimination was possible and the D.R. was about 50 per cent. In the case of the other two pairs of imposed stimuli, the discrimination improved gradually until it reached a high level.

FIGURE 2



Daily results in Experiment 2. Section (*a*) shows the performance on the imposed discrimination as a D.R. Section (*b*) shows the number of switches made to the alternative stimuli. The stages are marked at the top of the figure. In stage A the imposed stimulus was white; in stage B the blue/yellow discrimination was present and in stage C the cross/circle discrimination was present.

In section (*b*) the number of switches made each day is plotted. When the blue-yellow discrimination replaced the white imposed stimulus, there was an initial rise in the number of switches in the case of B2 and a subsequent gradual fall to a final low value. B3 showed no initial increase and gradually reduced the number of

switches he made. B4 showed a small rise on the first day of the blue-yellow discrimination and then a small drop on the next 3 days. At this stage he was still meeting high ratios on the bar especially when the imposed stimulus was blue, and the ratio requirement was increased to 33 for nine sessions, which are omitted from Figure 2 at the point marked by the arrow. Eventually the number of switches fell to a value of about six and most of these occurred when the blue stimulus was present, to which he may have had an aversion.

When the cross-circle discrimination was presented, all the animals showed a small increase in the number of switches which was followed by a fall, and when the white stimulus was reintroduced, the number of switches increased once more.

Discussion

The result of this experiment is in keeping with that of Experiment 1. The preference for the alternative stimuli fell as the animals became familiar with the imposed blue-yellow discrimination, and rose again when the new cross-circle discrimination was presented. This later rise was small but all the animals learned the discrimination quickly.

The data are in keeping with the hypothesis that the reinforcing strength of the alternative stimuli increases with the amount of uncertainty they reduce and suggest that this holds when uncertainty is varied in different ways. These animals switched from the white imposed stimulus even though it was the first time they had come across it, and it was not the indiscriminable end point of a discrimination as it had been in Experiment 1. This suggests that there is nothing specific about difficult discriminations in this context, beyond the fact that they are associated with uncertainty.

Time did not permit the study of a wider range of discriminations, but it would be interesting to present more difficult discriminations which would take longer to be learned. Another interesting experiment would be to study the formation of learning sets with this technique (Harlow, 1949). It is possible that with experience on a large number of discriminations, the preference for the alternative stimuli would not increase as a new discrimination was presented. One aspect of learning set formation might be that stimuli become reinforcing even before a discrimination is established so that observing responses are strong.

Experiment 3

It was pointed out earlier that with the present experimental arrangement, the animals could make use of the information provided by the alternative stimuli. When discriminative stimuli were absent in the imposed condition, the animals panel pressed on all trials, while in the alternative condition they were able to pause on negative trials. The overall response cost per reinforcement was, therefore, smaller in the alternative condition, and it is possible that the reinforcing strength of the alternative stimuli depended on this saving in panel presses.

If this were true, the reinforcing strength of the alternative stimuli should disappear if the response requirement on the panel was eliminated and the procedure converted to one involving classical conditioning where no specific response was required to produce food. The purpose of the present experiment was to observe the changes in the preference for the alternative stimuli as the operant procedure used so far was converted to a classical procedure. Reinforcement was, therefore, delivered at the end of positive trials and withheld on negative trials without any specific response being required. Had the microswitch behind the panel been simply disconnected, it would probably have taken a long time for panel pressing to extinguish, and some pressing may have been maintained indefinitely by superstitious association with reinforcement (Skinner, 1948). In the case of B1, therefore, the panel was locked so that it could not be moved. Surprisingly this produced virtually no change in his behaviour and he continued to press the locked panel,

even though it did not move. These presses were not counted, since the microswitch behind the panel was not being operated, but from general observation the rate and pattern of panel pressing remained unchanged. The panel was, therefore, unlocked and panel pressing discouraged by arranging each press to delay the end of the trial. At first short delays were used and these were gradually lengthened to 15 sec. In view of this experience with B1 the other animals were placed directly on the delay situation.

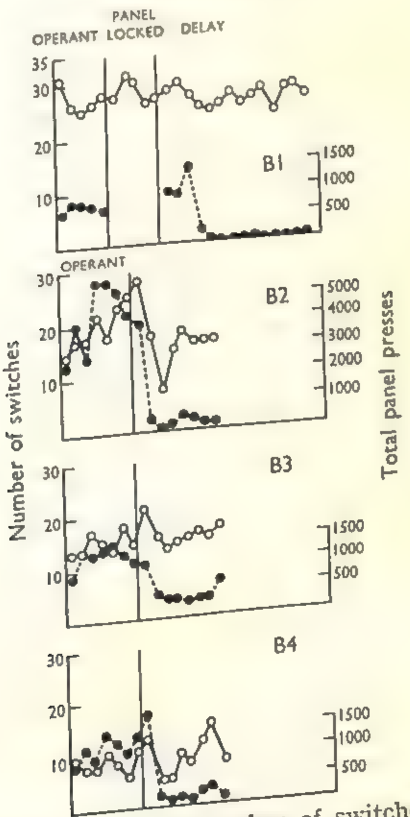
In other respects the procedure was the same as that used in the previous experiment. Positive and negative trials occurred in the same sequence, and bar pressing, as before, resulted in the presentation of a red stimulus on positive trials and a green one on negative trials, the imposed stimulus being white on all trials. After 65 sec. had elapsed the trials terminated and a reinforcement was delivered on positive trials. If, however, the animal pressed the panel, the termination of the trial was delayed. The length of the delay was increased over a few days to 15 sec. No change was made in the method of switching to the alternative stimuli.

Four animals, B1, B2, B3 and B4 were used as subjects, and the step by which the ratio increased was two in the case of B1, three for B2, and 10 for both B3 and B4.

Results

The circles in Figure 3 show the total number of panel presses made in each session, and it can be seen that with the introduction of the delay these fell. B1

FIGURE 3



Daily results in Experiment 3. The number of switches made to the alternative stimuli are shown as open circles. The filled circles show the total number of panel presses made in each session. In the case of B1 the panel was first locked for five sessions, and then each panel press was arranged to delay the termination of the trial for 1 sec. on the first session, 2 sec. on the second session, 10 sec. on the third, and 15 sec. on subsequent sessions. In the case of the other animals the panel was not locked and the following delays were instituted at the first vertical line; 5 sec. on the first session, 10 sec. on the second, third and fourth sessions, and 15 sec. on subsequent sessions.

stopped panel pressing completely, but the other animals continued to press at the beginning of some trials. This panel pressing was almost entirely restricted to those trials on which the animals bar pressed, and was probably maintained by a superstitious association between panel pressing and a switch to the alternative stimuli. The animals would press the bar and the panel at the same time, and once the alternative stimuli appeared, would stop pressing altogether. On those trials when they did not bar press, they also did not panel press. Panel pressing could probably have been eliminated completely by introducing a delay between each panel press and the appearance of the alternative stimuli, but this was not thought necessary since the panel pressing was clearly not related to food reinforcement and the animals did not press towards the end of the trial.

It can be seen that despite these changes in panel pressing, there was very little change in the preference for the alternative stimuli. A temporary drop in the number of switches occurred in the case of B2, who had been panel pressing at an extraordinarily high rate before the delay was introduced, but virtually no change was seen in the other animals.

Discussion

The failure to find any change in the number of switches made to the alternative stimuli as panel pressing decreased, shows that the reinforcing strength of the alternative stimuli did not depend on the saving in panel presses which the information allowed. It is possible, of course, that some alternative response, such as the superstitious behaviour which was seen in two of the animals, could play the same functional role as panel pressing and account for the results. Although it is not possible to exclude this, it seems unlikely that such responses could have developed so quickly and consistently in all of the animals, or that the saving in work should have been precisely the same as when panel presses were involved.

An alternative possibility is that the cost in terms of panel pressing was relatively trivial, and that the usefulness of the information was principally in terms of such preparatory autonomic responses as salivation. The informative stimuli could result in a saving of unnecessary drops of saliva, etc., and as Perkins (1955) and Prokasy (1956) suggest, enable preparatory responses to be appropriate to the type of trial and hence enhance the value of the primary reinforcement.

Finally it is possible as Berlyne (1960) suggests, that uncertainty leads to conflict, perhaps between competing autonomic responses or even "expectancies" in the classical situation, and that informative stimuli are reinforcing because they reduce this conflict.

These hypotheses would require elaboration but they could be distinguished experimentally. For example conflict might give rise to anxiety, the autonomic manifestations of which could be recorded, while the preparatory value of salivation would be less if wet food was used as reinforcement. The data from these experiments, therefore, confirm the fact that informative stimuli are reinforcing, and show that their reinforcement strength increases with the amount of information provided. They do not establish why the stimuli are reinforcing, but show that it is not simply due to the saving of responses.

This investigation was supported by the Medical Research Council. I would particularly like to thank Dr. L. Weiskrantz for his helpful advice.

References

Attneave, F. (1959). *Application of Information Theory to Psychology*. New York: Holt.

BERLYNE, D. E. (1960). *Conflict, Arousal and Curiosity.* New York: McGraw Hill.
BROADHURST, P. L. (1960). Abnormal animal behaviour. Ch. 19 in EYSENCK, H. J. (Ed.). *Handbook of Abnormal Psychology.* London: Pitman.
GARNER, W. R. (1962). *Uncertainty and Structure as Psychological Concepts.* New York: Wiley.
HARLOW, H. F. (1949). The formation of learning sets. *Psychol. Rev.*, **56**, 51–65.
HODOS, W. (1961). Progressive ratio as a measure of reward strength. *Science*, **134**, 943–4.
KELLEHER, R. T. (1958). Stimulus producing responses in chimpanzees. *J. exp. Anal. Behav.*, **1**, 87–102.
KELLEHER, R. T., RIDDLE, W. C., and COOK, I. (1962). Observing responses in pigeons. *J. exp. Anal. Behav.*, **5**, 3–13.
PAVLOV, I. P. (1927). *Conditioned Reflexes.* London: Oxford Univ. Press.
PERKINS, C. C. (1955). The stimulus conditions which follow learned responses. *Psychol. Rev.*, **62**, 341–8.
PROKASY, W. F. (1956). The acquisition of observing responses in the absence of differential external reinforcement. *J. comp. physiol. Psychol.*, **49**, 131–4.
SKINNER, B. F. (1948). Superstition in the pigeon. *J. exp. Psychol.*, **38**, 168–72.
STEINER, J. (1964). *The Reinforcing Properties of Discriminative Stimuli.* Ph.D. Dissertation. University of Cambridge.
WYCKOFF, L. B. (1952). The role of observing responses in discrimination learning: Part I. *Psychol. Rev.*, **59**, 431–42.
ZEIGLER, H. P., and WYCKOFF, L. B. (1961). Observing responses and discrimination learning. *Quart. J. exp. Psychol.*, **13**, 129–40.

Manuscript received 15th August, 1966.

CLASSIFICATION ON THE BASIS OF CONDITIONAL CUES

BY

R. DAVIS and D. H. TAYLOR

From the Department of Psychology, University of Reading

Subjects were presented with signals from two sources, either simultaneously or successively. The signal from one source, the "cue" signal, provided information about the type of classification to be made; the signal from the other source provided the data for the classification. Reaction times were recorded from the moment of presentation of the second signal to the moment the subject pressed the appropriate response key. By varying the order of presentation and the time interval between cue and data signal, the time required for the subject to select the appropriate dimension of variation in the data signal was examined.

When the cue signal preceded the data signal, the results were consistent with consecutive sequential processing of the two signals. When the data signal preceded the cue signal a further source of delay was evident. It is suggested that this results from a type of intermittency in which processing of irrelevant aspects of the data signal holds up the analysis of the cue signal.

Introduction

In a previous paper (Davis, 1964) one of the authors examined a situation in which information from two different sources had to be combined in order to select an appropriate response. By varying the interval between the two signals and measuring the reaction time for the appropriate response from the moment of occurrence of the second signal, the time course of the decision procedure was examined.

In this situation both signals played a similar role. If the second signal had been exchanged with the first, the total situation would have remained unchanged. There are other situations in which information from different sources has to be combined in rather a different way. An example, of some practical importance, is a situation in which one of the signals varies along several perceptual dimensions and the relevance of any particular dimension for purposes of classification is determined by the nature of the other signal. That is, one signal, the cue signal, provides information about the type of classification to be made whereas the other signal, the data signal, provides the data for the classification. This might be regarded as a filtering situation in which the nature of the information to be filtered from the data signal is determined by the cue signal. Much has been written on filtering as a basis for selective attention (cf. Broadbent, 1958; Treisman, 1964) but in most situations to which it has been applied, the relevant aspects of the stimulus situation which the subject is required to filter have been determined well in advance of the presentation of the stimulus material (usually by the experimenter's instructions).

If one uses a technique in which the instructions (or cue signal) as to which dimension is relevant are presented simultaneously with the data to be classified, or at a controlled time beforehand, the precise way in which the selective instructions operate and the time taken to achieve the selection can be examined in detail. Leonard (1958) used a similar technique, in investigating partial advance information, to examine the effect of limiting choice within a single dimension (a set of six lights) by giving a "cue" signal which indicated which half of the set would contain the subsequent signal to respond.

The technique can be further extended to examine situations in which the data signal is presented before the cue signal. In this case the subject receives the

information he has to classify before receiving the instructions how to classify it, and several alternative strategies may be available to him.

The complete experiment therefore was designed to investigate situations in which cue and data signals are given (*a*) simultaneously, (*b*) cue signal followed by data signal, (*c*) data signal followed by cue signal, with a controlled time interval between.

Method

Apparatus

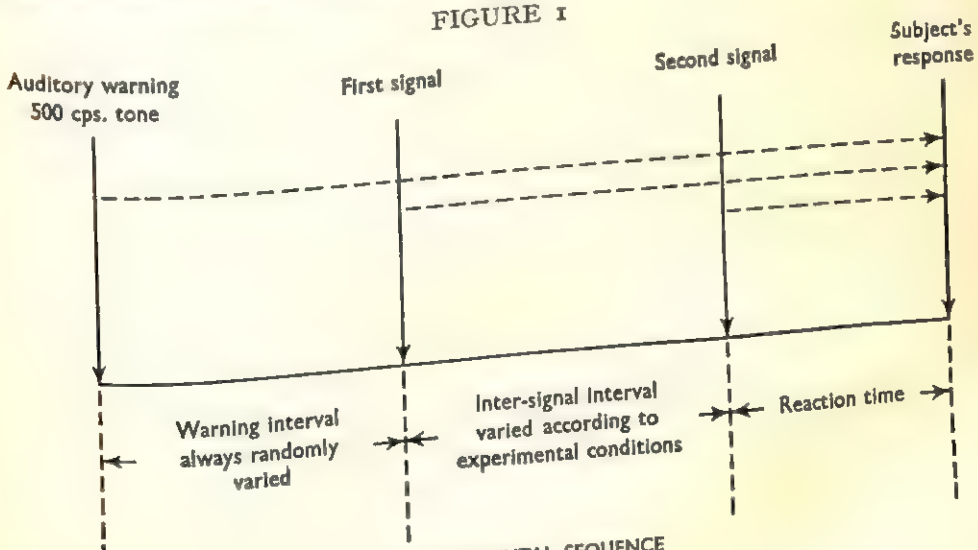
Signals were provided by two projection-type digital display units mounted one immediately above the other, about 3 ft. in front of the subject who was seated at a small desk in a sound insulated cubicle. The upper indicator provided the cue signal, which was either + or o. The data signal was provided by the lower indicator and consisted of a numeral, either 1 or 2, presented in the centre of a coloured disc, which was either red or green. It was arranged that signals reached 90 per cent. of maximum brightness within 5 millisec. of onset and both cue and data signals remained on until they were switched off by the subject's response.

The subject responded by pressing one of four microswitch pushbuttons, and these pushbuttons were so arranged that they could be comfortably operated by the first or middle finger of either hand.

The nature of the signals presented and the time intervals between them were controlled by the experimenter from outside the cubicle. Reaction times from the start of the second signal were recorded to the nearest 2 millisec. by a specially constructed apparatus, which produced punched paper tape, giving the details of signals and time intervals presented, as well as reaction times, in a form suitable for subsequent analysis by computer. The subject was not informed of his reaction times, but if he pressed an incorrect button (or more than one button) a buzz in a small loudspeaker sounded as an error signal. The same loudspeaker was used to provide an auditory warning signal, consisting of a 500 cps. tone which started either 768, 1,024, 1,280 or 1,536 millisec. before the onset of the first visual signal and continued until the subject made his response. The four values of warning interval were used randomly, with equal frequencies through all the conditions of the experiment.

The sequence of events is shown in Figure 1.

FIGURE 1



EXPERIMENTAL SEQUENCE

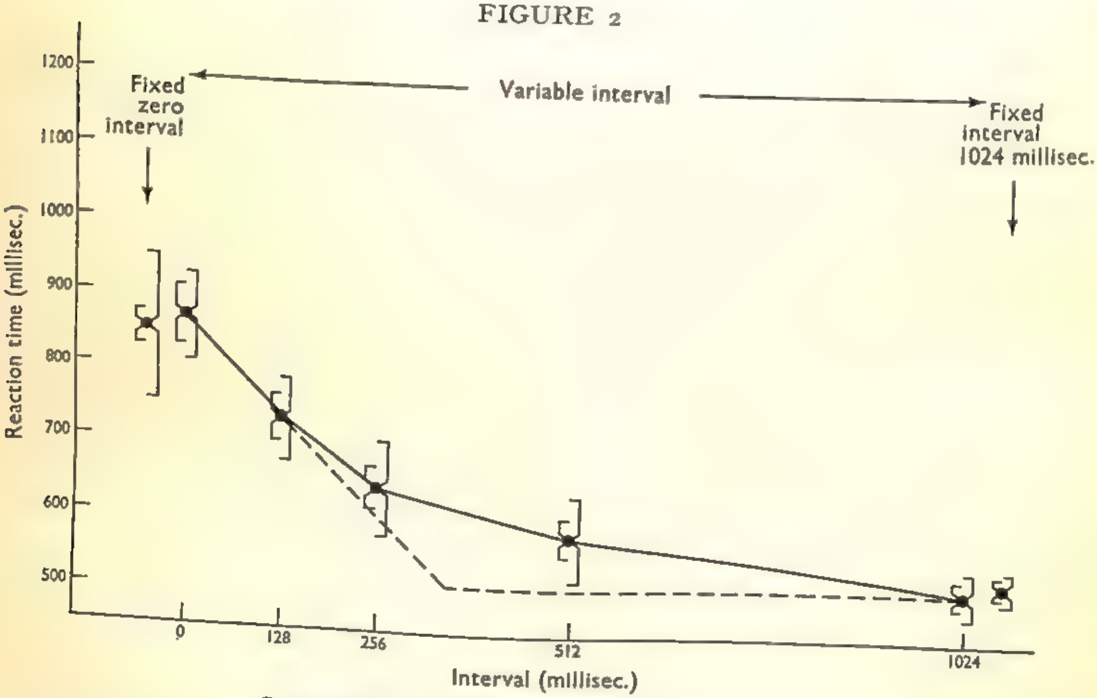
Procedure

Sixteen paid volunteers, five men and 11 women, between the ages of 18 and 26 acted as subjects. Each subject attended one training session of about ½ hr. and a few days later an experimental session lasting about 1 hr.

The subjects' task was to select on the basis of the cue signal, + or o, the appropriate stimulus dimension of the data signal. The cue signal "+" was always used to indicate that number was the relevant dimension and the cue signal "o" that colour was the relevant dimension. The stimulus dimension determined whether the correct response was to be made with the right or the left hand, and the value of the dimension, 1 or 2, in the case of the numbers, red or green in the case of colours, determined which finger of the appropriate hand was correct. In instructing the subject the experimenter tried to avoid suggesting that the classification should be made in any particular order. The eight possible stimulus-response configurations were balanced randomly among the 16 subjects and the configuration for any one subject remained the same throughout training and experimental sessions.

The training session consisted of a short introductory series of trials to familiarize the subject with the task, followed by a single run of 64 trials. During the training session cue and data signal were always presented simultaneously. After subjects had completed this session they were divided into two groups of eight, matched as nearly as possible on the basis of their average reaction times, over the run of 64 trials.

FIGURE 2



Condition C. Cue signal before data signal.

Mean reaction times over eight subjects for the variable interval run and the two fixed interval runs. The narrower confidence interval represents 90 per cent. confidence limits based on the average standard error for the individual reaction times for each subject. The broader confidence interval represents 90 per cent. confidence limits based on the standard error of the means over all subjects.

Each of the matched groups was given one of two treatments in the experimental session.

Condition C. For one group the cue signal appeared either at a controlled time *before* the data signal or simultaneously with it according to the schedule given below.

Condition D. For the other group the cue signal appeared either at a controlled time *after* the data signal or simultaneously with it according to the schedule below.

The experimental session consisted of three runs separated by short rest periods. In the first run the interstimulus interval was zero, as in the training session, and 64 trials were given.

In the second run of 160 trials the interstimulus interval was randomly varied using the following values with equal frequency. 0, 128, 256, 512, 1,024 millisec.

In the third run of 64 trials the interstimulus interval was held constant at 1,024 millisec.

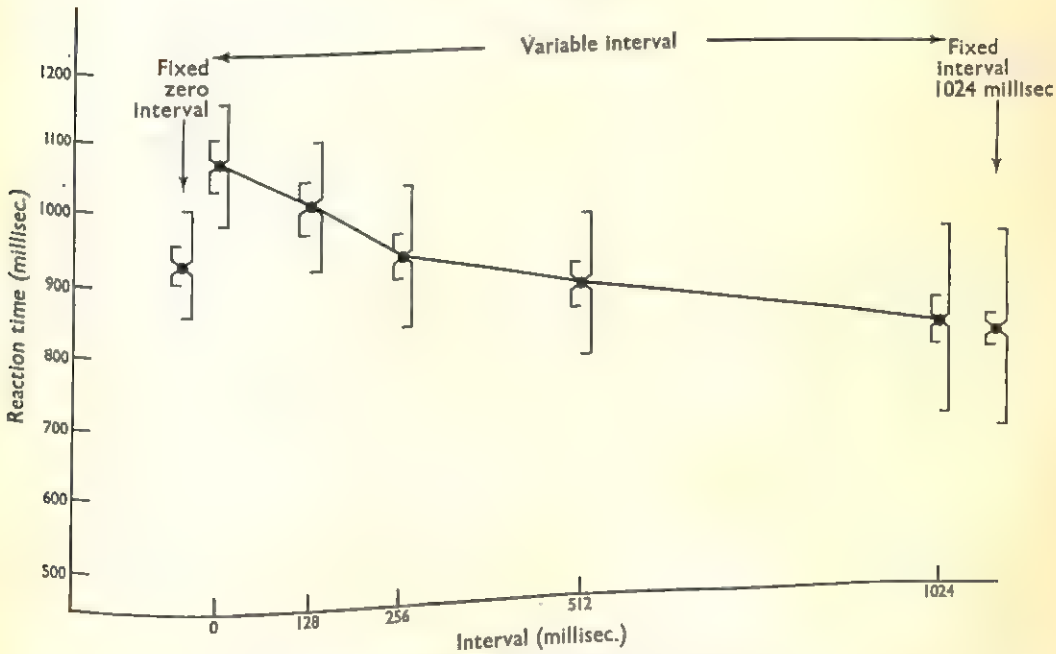
Within each run the different combinations of cue and data signal were randomized with equal frequency.

Results

Results are presented graphically in Figure 2 for Condition C and in Figure 3 for Condition D in the following form.

Mean reaction times for the group of eight subjects for each condition/interval in the experimental session are plotted, together with two confidence intervals. The narrower confidence interval represents 90 per cent. confidence limits based on the average standard error for the individual reaction times for each subject. It thus gives an estimate of within-subject variation. The broader confidence interval represents 90 per cent. confidence limits based on the standard error of the means over all subjects. It thus gives an estimate of between-subject, or population, variation.

FIGURE 3



Condition D. Data signal before cue signal.

Mean reaction times over eight subjects for the variable interval run and the two fixed interval runs. The narrower confidence interval represents 90 per cent. confidence limits based on the average standard error for the individual reaction times for each subject. The broader confidence interval represents 90 per cent. confidence limits based on the standard error of the means over all subjects.

The corresponding values for the 64 trials at zero fixed interval and for the 64 trials at 1,024 millsec. fixed interval are also shown respectively to the left and right of each graph.

Discussion

It is immediately apparent that subjects behave differently under the two conditions C and D.

Condition C. Cue signal followed by data signal. If subjects process information from the two signals sequentially, with processing of data information following immediately after processing of cue information, this strategy would be represented by a function with a linear slope of -45 degrees from its value at zero interval, reducing to a value equal to the processing time for the data signal at an interval equal to the processing time for the cue signal. This ideal strategy is represented by the dotted line on the graph, on the assumption that at 1,024 millsec. an asymptote corresponding to the processing time for the data signal has been reached. This assumption will be considered in more detail below. It can be seen that the actual results follow this pattern over the first 128 millisec. although the mean reaction times at 256 millisec. and at 512 millisec. are slightly longer than would be expected. It might be argued that temporal uncertainty of the second signal could contribute to a lengthening of reaction times over those expected in an ideal sequential strategy. However, for the two points at which a series of trials at fixed intervals were used, viz. zero and 1,024 millisec., the results for fixed and variable values of these intervals differ so little that temporal uncertainty cannot be contributing much to reaction times, at least at the extremes of the range of intervals used. The results therefore follow fairly closely the pattern that would be expected from sequential processing with non-overlap of the processing time for cue and data signal, although this leaves unaccounted for the lengthening of reaction times in the middle of the range of intervals used.

Further insight into the nature of the classifying operations in this condition may be gained by comparison with the results of a preliminary experiment, with four different subjects, which was used as a pilot for the present experiment and which was identical in all respects with the C condition, except that two further control conditions were used.

Control condition 1. In this condition the cue signal was presented simultaneously with the data signal but was not varied from trial to trial. The subject therefore knew in advance which was the relevant dimension of the data signal to respond to. This condition therefore represents a two choice situation along one dimension of the data signal with the other dimension always irrelevant and known to be so in advance. Analysis of variance of these results showed that there was no significant difference at $p < 0{\cdot}05$ for reaction times in this condition and for the variable condition at inter-signal interval of 1,024 millisec., in which the cue signal was changed randomly from trial to trial.

This control provides confirmation that at the longest interval used the processing of the cue signal is complete and that the only further processing required is a binary classification of the relevant dimension of the data signal.

Control condition 2. In this condition no cue signal was presented but at the end of the warning interval the data signal arrived and one of four responses had to be made. The value of one dimension of the data signal determined which hand was required and the value of the other dimension determined which finger of the hand was required for a correct response. This condition therefore represents a four choice situation with the stimuli having two sensory dimensions each determining a different feature of the response. Analysis of variance of these results established that there was no significant difference at $p < 0{\cdot}05$ for reaction times in this condition

and for the classification of cue signal presented simultaneously with the data signal as in the present experiment.

This control establishes the equivalence in terms of processing time between (A) classification of a signal along each of the two binary dimensions used for the data signal, and (B) selection of a relevant binary dimension followed by a binary classification along this dimension.

Further it might be observed that the operations described under (A) constitute a type of logical "AND" decision whereas the operations described under (B) constitute a type of logical "OR" decision.

Condition D. Data signal followed by cue signal. In this situation at least two alternative strategies may be distinguished.

(1) No processing of the data signal may occur until the cue signal arrives. In this case there should be no change in the value of reaction times as the interval between data and cue signal is varied—the same amount of processing always commencing with the arrival of the cue signal in each case.

(2) Processing of the data signal may start immediately and the classification could be reduced to a choice between two alternatives by the time the cue signal arrives, although in order to do this both dimensions of the data signal would have to be classified and the final choice would be across dimensions.

The actual results do not conform with a consistent application of either of these strategies.

The most striking point is that as the intersignal interval falls from 256 millsec. to zero, the reaction times for the variable interval show a progressively increasing delay (up to about 140 millsec.) with respect to the zero fixed interval value.

Furthermore the values of reaction times for the longest interval of 1,024 millsec. (both fixed and variable) show a reduction of the order of only 100 millsec. in relation to reaction times at zero fixed interval in contrast to a reduction of over 300 millisec. in Condition C. It therefore seems clear that if any processing of the data signal takes place before the arrival of the cue signal, this does not contribute greatly to a reduction in total processing time even at the longer intervals used. At the shorter intervals, less than 256 millisec., there appears to be some kind of interference leading to even longer reaction times than would be expected if processing commences with the arrival of the cue signal.

A possible explanation of this effect may be offered in terms of a "refractory period" similar to that found in other situations by the first author (e.g. Davis, 1957, 1959).

Thus one may suppose that as soon as the data signal arrives, the central systems become occupied in an attempt to classify it. However since there is no cue signal to indicate the relevant dimension, this classification may well start with the irrelevant dimension (on 50 per cent. of occasions if there is no bias towards one or other dimension) and this information is of no further use.

Furthermore, if the cue signal arrives while this analysis is taking place, then it would be consistent with previous work mentioned above that analysis of the cue signal would not proceed until the central systems were cleared of the previous analysis. This form of intermittency could thus give rise to delays in reaction time of the kind observed at the shorter intervals.

This work was undertaken as part of an investigation on the presentation of information in skill situations financed by the Science Research Council at the University of Reading.

References

Broadbent, D. E. (1958). *Perception and Communication.* London: Pergamon Press.

Davis, R. (1957). The human operator as a single channel information system. *Quart. J. exp. Psychol.*, **9**, 119–29.

Davis, R. (1959). The role of "attention" in the psychological refractory period. *Quart. J. exp. Psychol.*, **11**, 211–20.

Davis, R. (1964). The combination of information from different sources. *Quart. J. exp. Psychol.*, **16**, 332–9.

Leonard, J. A. (1958). Partial advance information in a choice reaction task. *Brit. J. Psychol.*, **49**, 89–96.

Treisman, A. M. (1964). Selective attention in man. *Brit. Med. Bull.*, *20*, **1**, 12–6.

Manuscript received 4th June, 1966.

ERROR-DETECTION AND CORRECTION LATENCIES AS A FUNCTION OF S-R COMPATIBILITY

BY

P. M. A. RABBITT* and SHIRLEY PHILLIPS†

From the Medical Research Council, Applied Psychology Research Unit, Cambridge and N.I.C.H.D., National Institutes of Health, Bethesda, Md., U.S.A.†*

In choice-response tasks employing correction-procedure, error-correcting responses are typically found to be faster than equivalent correct responses. An experiment was made to compare error-correction RT under conditions of good and poor S-R compatibility in a two-choice task. After practice, variations in S-R compatibility producing significant variations in mean correct RT nevertheless have no effect on error-correction time. The contrast between this result, and one previously reported (Burns, 1965) leads to a re-discussion of the processes of error detection and correction.

Introduction

In laboratory choice-response tasks and in many industrial situations human operators are required to correct errors by immediately making the response which they should have made. Two recent investigations have shown that such "error-correcting responses" may be faster than correct responses (Rabbitt, 1966) or even than correct response repetitions (Burns, 1965).

To explain this fact Burns (1965) and Rabbitt (1966) suggested that errors are not made at random, but are, as it were, approximations to correct responses. A subject who has committed an error thus, nevertheless, has partial advance information as to what the correct response should have been. In a task with correction procedure he can choose his correcting response relatively quickly from among a reduced number of alternatives.

In support of this hypothesis Burns describes the results of an eight-choice, paced, continuous-performance task with one-to-one Stimulus-response mapping. The response-stimulus interval was constant at either 150 millisec. (Condition 1) or at 820 millisec. (Condition 2). Under these R-S intervals he examined correction-time under two conditions of S-R compatibility: Direct-Correspondence (DC) where the neon-bulb signals mapped directly from left to right onto the response-keys and Mirror-Image (MI) where the right-hand keys mapped on to the left-hand lights, and vice-versa.

In the DC condition error-correcting responses were always faster than correct responses or than response-repetitions. The difference-score for Mean Correct — Mean-Error-correcting RT was significantly greater with a long (820 millisec.) than with a short (150 millisec.) R-S interval. In the MI condition error-correcting responses were again significantly faster than equivalent correct responses at the long R-S interval, but not at the short R-S interval. Overall, difference scores for Mean Correct — Mean Error-correcting RT were significantly greater in the DC than in the MI condition.

Burns argues that these results are consistent with the results of other experiments in which subjects performing choice-response tasks were given fore-signals which partially reduced their uncertainty as to which of a set of reaction-signals was to be presented. In such tasks it has also been shown that the reduction in RT after an informative fore-signal is a direct function of the interval between fore-signal and reaction-signal (Leonard, 1958). Recent theoretical statements may also

be taken to suggest that, for a given fore-period, this reduction should be greater with a compatible than with an incompatible S-R mapping (Lawrence, 1963). Since analogous variations in experimental conditions produce similar effects on the RTs of error-correcting responses, Burns concluded that error-correction is relatively fast because the subject's error gives him partial advance information as to what the correcting response should be.

The correction of an error must involve at least two processes which are logically distinct, though they may or may not take place concurrently: the error must be *detected* and then an appropriate response must be *selected* from among the remaining possibilities. Assuming that the subject takes some time to realise that he has made an error, we would expect Burns's finding that the RT of the error-correcting response is reduced as the response-stimulus interval is increased. Burns's elegant experiment also does not rule out the possibility that errors may take longer to *detect* with his MI than with his DC mapping. In this case we would also expect his finding that error-correcting responses are relatively slower in the MI condition. The hypothesis of partial information from errors will only account for Burns's data completely if the time taken to recognize an error is independent of S-R compatibility.

Burns's data also do not exclude the possibility that error-correcting responses are facilitated by other factors than the subject's possession of partial advance information. Burns himself preferred to suggest that the occurrence of an error was followed by an orientating response (Sokolov, 1963) which inhibited rather than facilitated subsequent responses. On this assumption the benefit of partial advance information would have to overcompensate for such inhibition. However, we may reasonably ask how the time taken to *detect* the commission of an error compares with the time taken to identify other, external, signals to which the subject responds.

These uncertainties as to the process of error-*detection* are weaknesses in Burns's argument. A task in which a subject selects between only two responses is a test case, for here the error-correcting response is always completely specified once an error is made (it is always the response which the subject did not make). In this situation, once a subject has made an error he can ignore the display (and so the conditions of S-R mapping) in order to make the alternative response. Since he does not have to pay any attention to the S–R mapping in order to correct an error, differences in the latencies of error-correcting responses with S-R compatibility must here imply differences in the time taken to detect that an error has occurred.

The experiment described below was made to test this.

Method

Apparatus. Subjects were tested on a "Psychomet" (cf. also Rabbitt, 1966). From the subject's point of view this apparatus consisted of a control panel set at an angle of 25° to the vertical and mounted on a desk at which he sat. A signal-lamp was mounted behind each of 10 circular apertures 1 in. wide and ½ in. apart from edge to edge, which were set in a horizontal row across the upper half of the panel. Translucent shielding diffused light so that when any lamp came on the aperture appeared as an evenly illuminated disc. Ten contact grids were mounted in a horizontal row across the bottom of the panel, each grid set 6 in. vertically below one of the apertures. These grids were 1 in. square and ½ in. apart.

Programmes read by a Rheem photoelectric reader and ancillary apparatus allowed the signal lamps to be switched on one at a time in any sequence desired. The response designated correct was also specified by this programme. A touch on the "correct" grid immediately switched off the lighted lamp, and the next lamp specified by the programme was lit within 20 millisec. The elapsed time between successive responses, whether correct or wrong, was measured to within 0·01 sec. by a version of SETAR (Welford, 1952). The SETAR output for each response specified whether the response was correct or wrong, and which grid was touched. Errors were timed and recorded and no auditory cue from

the apparatus discriminated between errors and correct responses. Nevertheless a comparator circuit ensured that the signal lamp did not change until an error was corrected by a touch on the "correct" grid specified by the Rheem programme. Thus subjects were always obliged to correct errors by making the response which they should have made before the experimental sequence of signals could continue. For this reason sequences were made up so that no signal was ever immediately repeated—the fast onset of signal lamps made it likely that an immediate repeat of the same signal would be confused with the failure of the display to change when an error occurred.

Procedure. Eight experimental sequences of 500 signals were programmed from a table of random decimal digits with minimal alteration necessary to ensure three constraints:

(1) No signal was repeated.

(2) All signals appeared equally often.

(3) Transitions between signals in the set of left-hand lamps 1 through 5 and signals in the set of right-hand lamps 6 through 10 occurred exactly as often as transitions between signals within either of these sets.

Four of these sequences were programmed for each of two conditions of S-R compatibility:

Direct mapping (DM). The response grid appropriate to any of the left-hand lamps 1 through 5 was grid 5 (the left-hand one of the two centre grids). Similarly, the response grid appropriate to the right-hand lamps 6 through 10 was grid 6 (the right-hand one of the two centre grids).

Cross-over mapping (CO). The S-R mapping used in the DM condition was reversed so that response grid 5 was appropriate to lights 6 through 10 and vice versa.

In both conditions subjects used both hands, responding on grid 5 with the left forefinger and on grid 6 with the right forefinger. It will be noted that due to constraint 3 (above) the number of times that responses were repeated was the same as the number of times that the subject alternated between responses.

One group of 12 subjects experienced four sequences on the DM condition while a second group of 12 experienced four sequences on the CO condition.

Subjects were employees at N.I.M.H. N.I.H. Bethesda, aged from 18·2 to 35·4 (Mean 26·3). None were aware of the purpose of the experiment.

Results

The output from SETAR was processed by an IBM. 1620 computer programmed to calculate mean response-time separately for correct response-repetitions and for correct alternations between grids. The computer programme was set to detect subject-errors and to ignore, in computing means, all errors and the first four responses following each error. A print-out of the entire data for each subject allowed spot-checks on the operation of the programme. Subject-errors were located by eye on this print-out, and mean response-times were tabulated for the response correcting each error and for the two responses following error-correction. In rare cases where errors occurred within three responses of each other, data from the earlier errors in such pairs or strings were separately tabulated. RTs of such responses were not found to differ significantly from latencies of "isolated" errors and all data were therefore pooled to give the results analysed below.

Incidence of errors. The incidence of errors in the DM condition varied between subjects from 1·1 to 6·8 per cent. (mean 3·2). Incidence in the CO condition varied from 2·6 to 9·2 per cent. (mean 4·8). An analysis of variance gave significant terms for differences in error-incidence between conditions ($p < 0{\cdot}001$). The interaction-term for practice × conditions was significant ($p < 0{\cdot}01$). Inspection of the data suggested that the significant interaction term occurred because practice reduced the incidence of errors more in the CO than in the DM condition. To check this, incidence of errors was compared between conditions by *t*-tests separately for the first two runs and for the last two runs in each condition. While the incidence of errors was

significantly greater in the CO condition early in practice ($p < 0.001$) there was no difference in the incidence of errors between the two conditions during the third and fourth practice runs.

RTs of correct responses. An analysis of variance on RTs of correct responses in each condition gave significant terms for differences between conditions ($p < 0.01$) and for practice ($p < 0.001$). The interaction-term for practice × conditions was also significant ($p < 0.01$). It appears that latencies are longer in the CO than in the DM condition, and that the reduction of RT with practice is relatively greater in the CO condition.

Error-correcting RTs. In Table I mean RTs of error-correcting responses are set out for comparison with RTs of correct alternations and repeats. To illustrate the effects of practice on each condition latencies for the four successive runs are separately tabulated.

TABLE I

MEAN CORRECT RT AND MEAN ERROR-CORRECTING RT (IN MILLISEC) FOR EACH OF 4 SUCCESSIVE RUNS OF 500 RESPONSES IN A 2 RESPONSE 8 STIMULUS TASK WITH DIRECT (DM) AND CROSSED-OVER (CO) STIMULUS-RESPONSE MAPPING

		Direct mapping condition						*Cross-over mapping condition*					
		Run 1	*Run 2*	*Run 3*	*Run 4*	*Repeats*	*Altern.*	*Run 1*	*Run 2*	*Run 3*	*Run 4*	*Repeats*	*Altern.*
Correct RT	Mean	502	479	482	464	450	504	672	546	523	515	502	681
	S.D.	58	45	43	50	41	69	74	65	68	58	52	84
Error-correcting RT	Mean	447	422	415	399			643	497	441	415		
	S.D.	73	54	55	49			112	74	69	60		

An analysis of variance was carried out on the data for each condition to compare error-correcting RTs with RTs of correct alternations and repeats. These analyses all gave significant terms for differences between these response-classes ($p < 0.001$). The error-term from each analysis was used to compute the S^2 statistic in order to rank-order the response-classes. For both conditions error-correcting responses were found to be significantly faster than either repetitions ($p < 0.01$) or alternations ($p < 0.001$). Repetitions were always faster than alternations ($p < 0.01$).

A further analysis of variance was made to compare latencies of error-correcting responses between conditions. Significant terms emerged for differences between conditions ($p < 0.01$), and for practice ($p < 0.001$). The conditions × practice interaction was significant ($p < 0.01$).

To further examine the implications of the significant interaction-term the error-term from this analysis of variance was again used to calculate S^2. It emerged that error-correcting responses were significantly slower in the CO than in the DM condition for the first ($p < 0.001$), second ($p < 0.01$) and third ($p < 0.01$) runs. However the difference between the fourth runs was not significant ($p > 0.1$).

It seems that at the beginning of practice latencies of error-correcting responses are significantly longer in the CO than in the DM condition. Nevertheless, after practice, latencies of error-correcting responses are the same in both conditions.

Difference Mean Correct RT—Mean Error-correcting RT

It was possible that the reduction of RT for error-correcting responses represented only one aspect of a general levelling-off of response-latencies between conditions (cf. latencies of correct responses, analysed above). In order to check this, difference scores for the values Mean Correct RT − Mean Error-correcting RT were obtained for

each subject in each run of each condition. Means of these means are given in Table II. An analysis of variance gave significant terms for differences between groups-conditions ($p < 0{\cdot}01$) and for order ($p < 0{\cdot}01$). The Conditions × Practice interaction was significant ($p < 0{\cdot}01$). S^2 was calculated from the error-term and was used to test differences between conditions at each level of practice in turn. It appeared that in runs 1, 2 and 3 difference scores for Mean Correct — Mean Error-correcting RTs were much the same in either condition ($p < 0{\cdot}1$) while difference

TABLE II

MEANS OF DIFFERENCE SCORES (IN MILLISEC) FOR MEAN CORRECT RT — MEAN ERROR CORRECTING RT FOR INDIVIDUAL SUBJECTS.

		Run 1	*Run 2*	*Run 3*	*Run 4*
Direct mapping condition ..	Mean	53	57	68	64
	S.D.	18	23	18	28
Cross-over mapping condition	Mean	24	39	72	129
	S.D.	19	25	34	48

(N.B.—These figures cannot be directly derived from Table I where overall means for correct and error-correcting responses are presented)

scores were greater in the CO than in the DM condition in run 4 ($p > 0{\cdot}01$). We may conclude that response-latencies in both conditions are reduced by practice, and that this reduction is more marked in the CO than in the DM condition. Further, in the CO condition reduction in the latencies of error-correcting responses is relatively greater than reduction in the latencies of correct alternations and repeats. Consequently, after practice, there is no significant difference between error-correcting RTs in the CO and DM conditions, and the difference Mean Correct — Mean Error-correcting is significantly greater in the CO than in the DM condition.

DISCUSSION

In this experiment the detection of an error also implied complete information as to which correction response should be made. After practice, conditions of S-R compatibility which significantly affected mean correct response-time nevertheless did not affect the time taken to correct errors. It follows that there is no reason to believe that the time subjects took to realize that an error had occurred was different for the two S-R mappings which were compared.

Burns (1965) proposed that error-correcting responses were relatively fast because the occurrence of an error implied partial advance information as to what correction response the subject had to make. In Burns's (1965) or Rabbitt's (1966) experiment such partial advance information might imply direction of the subject's attention to the relevant sector of the display, or (Burns, 1965) to the correct one of his two hands. This might indeed be expected to facilitate error correction when the correction response must be selected from among 7 (Burns) or 9 (Rabbitt) remaining possibilities. Nevertheless, Burns's hypothesis only completely explains his data on the assumption that subjects take no longer to detect that they have made an error when the S-R compatibility is poor than when it is good. To the extent that one can extrapolate from the relatively simple 2R/8S mappings used in this experiment to the complex 8R/8S mappings used by Burns, these data therefore support Burns's hypothesis.

Crucial, however, is the distinction that S-R compatibility affects latencies of error correcting responses only when the correction response must be selected from among alternatives (that is, with reference to the display, and so to the prevailing S-R mapping). The present experiment makes the point that the selection of an error-correcting response and the detection of an error are operationally separable processes, and that further investigations of error-correction must take this into account. Two characteristics of error-*detection* are so far evident, (1) The time taken to detect the occurrence of an error is independent of S-R compatibility within the conditions examined. (2) Subjects detect that an error has occurred much more quickly than they respond to any other signal presented in either condition (i.e. whether to signals requiring repeated or alternated responses). The qualifying riders to these statements are:

(1) These data, and all results so far published, show fast error-correction where subjects correct by making the response which they should have made. Data from tasks in which subjects make a common correction response to all errors are crucial to discussions of error-detection, but have not yet been reported.

(2) In two reported studies (Burns, 1965; Rabbitt, 1966) and in the present experiment, latencies for error-correcting responses were timed from the moment of subject's previous (wrong) responses. There is, of course, no guarantee that the process of error-detection was begun by the subject at this moment. A possible explanation of fast error-correction times is that a subject realises that he is about to commit an error before he completes his wrong response—thus overlapping the processes of error-commission and error-correction (cf. Rabbitt, 1966).

Since it is not likely that the amount of this overlap would be greater for non-compatible situations, the main part of the argument of this paper is not affected by these considerations. Nevertheless, we have as yet no grounds to believe that we know the *absolute* latencies for the detection and correction of an error.

The authors are glad to acknowledge use of the facilities of the Section on Aging, N.I.M.H., and the personal support of Dr. J. E. Birren.

References

Burns, J. T. (1965). *The Effect of Errors on Reaction Time in a Serial Reaction Task.* Unpub. Ph.D. dissertation the University of Michigan.

Lawrence, D. H. (1963). The nature of stimulus: Some relationships between learning and perception. In Koch, S. (Ed.) *Psychology: Study of a Science.* New York: McGraw Hill, **5**, 179–212.

Leonard, J. A. (1958). Partial advance information in a choice-reaction task. *J. gen. Psychol.*, **49**, 89–96.

Rabbitt, P. M. A. (1966). Errors and error-correction in choice-response tasks. *J. exp. Psychol.*, **71**, 264–72.

Sokolov, E. N. (1963). Higher nervous functions: the orienting reflex. In Hall, E. (Ed.) *Annual Review of Physiology.* Palo Alto: Annual Reviews Inc., **25**, 545–80.

Welford, N. (1952). An electronic digital recording machine—the SETAR. *J. Scient. Instrum.*, **29**, 1–4.

Manuscript received 22nd August, 1966.

THE DISTANCE GRADIENT IN THE AFTER-EFFECT OF KINAESTHETIC WIDTH JUDGEMENT

BY

A. W. PRESSEY

From the Department of Psychology, University of Manitoba

The combined effects of inspection-time and interfigural distance on a kinaesthetic figural after-effect were determined. The figural after-effect was defined as the degree to which a 2-in. width appeared to shrink following prolonged inspection of larger widths. The inspection-widths were 2·0, 2·5, 3·0, 3·5, 4·0, and 4·5 in. and the inspection-times were 10 and 50 sec. A control group which did not inspect any widths was also employed. 156 subjects were tested. The results indicated that the classical non-monotonic relationship between interfigural distance and figural after-effect was not present. Instead, the figural after-effect increased as a negatively accelerated function of interfigural distance. Increasing inspection-time increased the asymptotic level of the figural after-effect. These results were interpreted in terms of the effects of anchors on judgements of magnitude.

Introduction

In 1947 Köhler and Dinnerstein demonstrated that kinaesthetic judgements of a standard width can be distorted by prior experience with a different width. Their procedure consisted of rubbing a block of a certain width (I-width) and then judging the width of either a larger or a smaller block (T-width). They found that narrower T-widths appeared to shrink and that wider T-widths appeared to expand. The term "figural after-effect (FAE)" was applied to these perceptual changes since they seemed to be analogous to distortions found in vision (Gibson, 1933; Köhler and Wallach, 1944).

Subsequent investigations of kinaesthetic FAEs were concerned primarily with specifying the effects of spatial and temporal variables of the degree of distortion. One variable that has been found to affect the FAE is the size discrepancy between I- and T-widths (I-T distance). Charles and Duncan (1959) and Krauskopf and Engen (1960) reported that there was an optimum I-T distance which yielded a large FAE. That is, when the I- and T-widths were nearly equal the FAE was small, but as I-T distance increased, the FAE increased at first and then decreased. Both Charles and Duncan and Krauskopf and Engen contended that this non-monotonic relationship between I-T distance and kinaesthetic FAE was formally similar to the distance gradient present in visual FAEs.

Although the existence of a distance gradient in kinaesthesis seems adequately established, there is no evidence on the question of whether the gradient *varies* with changes in stimulating conditions. For example, do variables such as inspection-(I-) time or time after inspection—variables which are known to have a profound influence on the FAE—alter the gradient? And if the gradient is altered, what is the precise nature of this change? It was to this general problem that the present experiment was directed. The main objective was to measure a kinaesthetic FAE as a joint function of I-T distance and I-time.

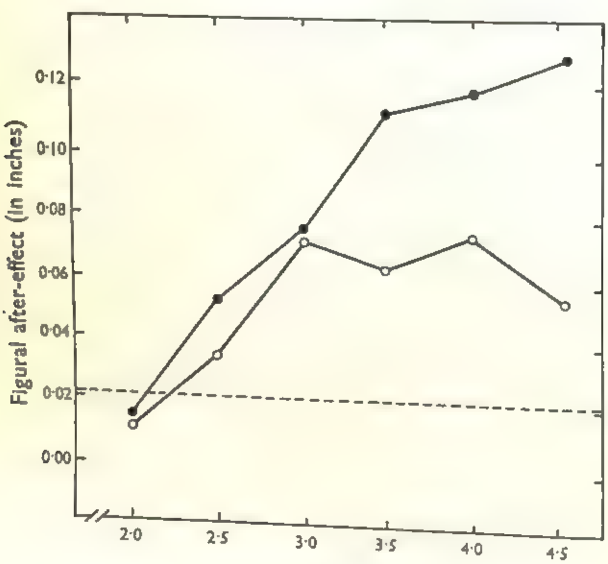
Method

Apparatus. The apparatus consisted of a comparator and 6 blocks of wood which varied in width from 2·0 to 4·5 in. in steps of 0·5 in. All blocks were used as I-figures but the 2-in. block was also used as the T-figure. Each block was 8 in. long and 2 in. high. The sides of the blocks were lined with 8- × 2-in. sheets of brass.

The comparator, described previously by Howarth (1964), was an 8- × 6- × 4·5-in. box which housed a reversible motor. Two brass bars, each 8 in. long and 2 in. high, were mounted parallel to each other on top of the box. One bar was stationary but the other could be moved toward or away from the stationary bar by means of the reversible motor. Direction of movement was controlled either by the subject with a foot pedal or by the experimenter with two hand switches. Depressing the pedal to the subject's left caused the distance between bars to decrease while depressing the pedal to the right increased the distance between the bars. The variable bar moved at a rate of 0·05 in. per sec. The distance between bars was measured accurately to within 0·001 in. by means of two micrometers located on the sides of the comparator.

Two black wooden stands were placed on either side of a chair. The stand on the subject's left was 7 in. high and the one on the right was 11·5 in. high. The seat of the chair was 18 in. above the floor. The comparator was placed on the left side and was oriented so that the stationary bar was 8 in. away from the chair. The I- and T-blocks were always located on the right side at a distance of 8 in. from the chair. With this arrangement the wooden blocks and the parallel bars were equidistant from the centre of the seat.

FIGURE 1



Kinaesthetic figural after-effect as a function of inspection-time and width of inspection block. The open circles represent 10 sec. of inspection and the filled circles represent 50 sec. of inspection. The broken, horizontal line represents the point of subjective equality obtained from the control group.

Subjects. The subjects were 156 students enrolled in introductory psychology at the University of Alberta, Edmonton. All of them had been tested on a visual FAE task just prior to the present study but none of them was familiar with the phenomenon of FAE.

Design. A 6 × 2 factorial design was employed. There were six levels of I-T distance which were obtained by using blocks ranging in width from 2·0 to 4·5 in. The T-block was always 2 in. wide. The two I-times were 10 and 50 sec. A control condition in which an I-block was not presented was also employed. The subjects were assigned at random to each of the 13 conditions in the order of their appearance at the laboratory.

Procedure. Each subject was told that his task was to make comparative judgements of width with the thumb and middle finger of each hand. The experimenter showed him the T-block and demonstrated how the comparator bars could be made equal to the width of that block. He was told that, in making his judgements, it was best not to

exert more pressure on the block or bars than was necessary to "just obtain a feel of the width." He was also asked to be as accurate as possible but not to spend too much time on each judgement.

The subject was blindfolded with opaque goggles and seated in the chair. Four trials were given to obtain a measure of the PSE. These trials consisted of two ascending and two descending adjustments given either in the order ADAD or DADA. Within each condition this order was alternated between subjects. On each trial the variable comparator bar was offset at a distance of 1·5 or 2·5 in. depending upon whether an ascending or a descending adjustment was to be employed. The subject grasped the T-block and the comparator bars and, upon a signal of "Depress right" or "Depress left" from the experimenter, set the bars equal to the width of the block. When the adjustment was completed, the subject placed both hands on his lap and relaxed. The time taken to make an adjustment by most subjects was between 10 and 20 sec. If a subject was still making a judgement after 15 sec., the experimenter prompted a decision by asking, "Is that about it?" The time interval between trials was 30 sec.

After the fourth trial a metronome was activated at a rate of one beat per sec. The subject was told that he was to grasp a block of wood between the thumb and middle finger and stroke the block back and forth in time to the metronome. The experimenter placed the subject's hand on the T-block and moved it back and forth for a period of 4 sec. in order to demonstrate the task. Direction of motion was reversed at every beat. When the subject understood the task, an I-block was substituted for the T-block and he rubbed that block for a period of either 10 or 50 sec. At the end of the I-period the metronome was turned off, the subject placed his hand on his lap, and the I-block was replaced by the T-block. The time required to exchange blocks was about 2 sec. The subject then made two ascending and two descending adjustments which were identical in all respects to the trials given prior to inspection.

The subjects in the control group were not asked to rub an I-block after the fourth trial. Instead they rested for 90 sec. before continuing with their adjustments.

Results

For each subject the four adjustments made prior to inspection were averaged to establish the PSE. A measure of the FAE was obtained by averaging the adjustments made after inspection and subtracting this average from the PSE. This resulted in a positive score for judgements which were narrower after inspection.

An analysis of variance was performed on the FAE scores. The effects of I-time and I-width were both statistically significant. The F-ratio for I-time was 4·77 ($F_{0\cdot05} = 3\cdot91$) and for I-width F was 3·44 ($F_{0\cdot01} = 3\cdot16$). The interaction between I-time and I-width was not significant ($F = 0\cdot70$; $F_{0\cdot05} = 2\cdot27$).

Changes in the FAE as a joint function of I-time and I-width are shown in Figure 1. It is clear that the FAE increases as a negatively accelerated function of the width of the I-block. Increasing I-time serves only to increase the *asymptotic level* of the FAE. That is, variations in I-time have a small effect when the difference between I- and T-widths is small but, as this difference increases, the differential effect of I-time also increases.

The scores from the control group were analysed by subtracting the mean of the last four adjustments from the mean of the first four adjustments. These difference scores were subjected to a t-test which yielded a value of 1·18. This result indicates that there was no significant shift in the PSE as a function of repeated trials.

Finally, an analysis was performed to determine whether there was any shift in the subjective width of the 2-in. T-block following prolonged stimulation with the 2-in. I-block. The FAE scores from the 10- and 50-sec. inspection groups were combined and a t-test for correlated measures was calculated. The resulting value of t was 0·81 which indicates that a kinaesthetic FAE does not occur when I- and T-figures are of the same width.

Discussion

The most surprising finding in this experiment is the fact that a distance gradient was *not* present. Instead, the FAE increased as a negatively accelerated function of I-width. There was no tendency for the FAE to decrease when very large I-widths were employed. This result is very similar to that obtained by Nachmias (1953). He measured changes in kinaesthetic judgement of a T-position following prolonged experience with different I-positions. A distance gradient was not found; the magnitude of displacement increased steadily with increasing I-T distance. Nachmias contended that a distance gradient might have been manifested had larger I–T distances been employed, but such an argument is not applicable to the present experiment. The largest I-width that was used was 4·5 in. which is very near the maximum span between the thumb and middle finger. Consequently, the evidence strongly favours the view that the decreasing portion of the distance gradient does not occur in after-effects of kinaesthetic width judgements.

The absence of a distance gradient in the present study immediately raises questions regarding the basis upon which Charles and Duncan (1959) and Krauskopf and Engen (1960) concluded that such gradients are present. Since these investigators also used kinaesthetic width judgement as their dependent variable, it seems unlikely that the contradictory conclusion lies in the type of sensory dimension that was distorted. A close examination of Charles and Duncan's and Krauskopf and Engen's experiments suggests that, in the former case, measurement of the FAE was confounded while the latter investigators misinterpreted their results.

Charles and Duncan used a single I-block which was 2 in. wide and several T-blocks which varied in width from 0·125 to 2·0 in. The I-block was presented for 60 sec. Their results showed that the amount of shrinkage of the T-blocks increased at first and then decreased as the width of the T-blocks decreased. But the procedure of varying T-block instead of I-blocks to increase I–T-distance is subject to criticism. Small T-blocks will not shrink as much as large ones regardless of the distorting conditions because different "ceilings" are involved. For example, whereas a 2-in. width may shrink more than 0·25 in., a 0·125- in. T-block cannot possibly shrink more than 0·125 in. This difference alone could cause FAEs to decrease when very narrow T-blocks are used. Charles and Duncan could have overcome this difficulty, at least partly, by taking the relative instead of the absolute amount of shrinkage as their measure of the size of the FAE. Their scores would then be expressed as per cent. decrement in width. When such a transformation is performed, it is difficult to argue that a distance gradient is present. A decrease in the FAE occurs only at the smallest T-width.

Krauskopf and Engen employed a 2-in. T-block and five I-blocks which ranged in width from 2·5 to 4·5 in. The largest FAE was present when the I-block was 4·0 in. and the FAE decreased when the 4·5 in. I-block was used. They interpreted this final decrement to mean that a distance gradient, formally similar to the one found in vision, is also present in the kinaesthetic FAE. But Krauskopf and Engen's conclusion is unconvincing because they failed to find a gradual increase in the FAE as I-widths increased from 2·5 to 3·0 to 3·5 in. Moreover, the largest (4·5-in.) block yielded a much greater FAE than the first three I-blocks. Hence, their data appear to fit a linear trend just as well, or better than, a quadratic one. Unfortunately, a trend analysis was not performed so that agreement or disagreement with Krauskopf and Engen's conclusion must be based on visual inspection rather than on statistical probabilities.

The failure to find a distance gradient in this study provides strong evidence against theories which utilize constructs such as satiation to explain kinaesthetic

FAEs (Köhler and Wallach, 1944). Implicit in the concept of satiation is the view that there is a definite limit to the extent to which satiation spreads beyond the region of maximum satiation. Regions well beyond the point of excitation are assumed to remain relatively unaffected. If the formal structure of satiation theory is to be applied to kinaesthesis, it would have to be argued that satiation in the kinaesthetic cortex must also be limited and that, therefore, a distance gradient should occur.

At the present time, the most promising approach to kinaesthetic FAEs appears to lie in frame-of-reference theories such as Helson's (1964) adaptation-level theory. Helson states that judgements of magnitude are always made in relation to a momentary indifference point which is termed "adaptation-level." Adaptation-levels increase whenever strong stimuli (anchors) are present and decrease whenever weak anchors are presented. The fact that a narrow width appears narrower after stimulation from a larger width is because the large width shifts adaptation level upward. When the narrow width is re-introduced, it is judged in relation to the new adaptation level and thus is perceived as narrower than it was prior to stimulation.

It cannot be denied that some simple facts, such as direction of distortion of kinaesthetic judgements, can be subsumed under Helson's theory. However, the key question that is prompted by the present experiment is whether extremely large anchors have the same effect as intermediate anchors or whether adaptation level is relatively unaffected by extremely large anchors. To put it more simply, the question is whether adaptation level theory predicts a monotonic or non-monotonic relationship between I-T distance and kinaesthetic FAE. Helson himself is not clear on this point because he states that an anchor may be so very different from the existing adaptation level that it may fail to affect that level. But this statement lacks theoretical import since it does not provide an independent operation by means of which one can establish when an anchor will be effective and when it will not be effective. Consequently, Helson's theory, as it now stands, does not allow an unambiguous prediction to be made regarding the relationship between I-T distance and the kinaesthetic FAE.

In contrast to the theoretical vagueness, the empirical data on the effects of increasing size of an anchor on magnitude of judgemental shifts are in substantial agreement. It has been known for a long time that categorical judgements shift systematically whenever an extreme stimulus is introduced into a series. Guilford (1954) summarized the work of several investigators who were concerned with the relationship between magnitude of anchoring stimuli on degree of judgemental shift and concluded that:

(a) The farther the anchor stimulus from the stimulus range, the greater the shift.

(b) The more remote the anchor stimulus, the less the *increment* of shift. By this is meant that the shifting effect shows diminishing returns.

It is evident that these statements describe precisely the relationship between size of I-block and magnitude of kinaesthetic FAE found in this study. The hypothesis that kinaesthetic after-effects result from a change in a frame-of-reference is, therefore, supported.

A final point may be noted. The view that kinaesthetic after-effects are produced by a shift in a frame-of-reference challenges the use of kinaesthetic after-effects to test neurophysiological hypotheses about satiation and reactive inhibition. There are several examples of such tests in the literature. Klein and Krech (1952) argued that larger kinaesthetic after-effects in brain-damaged patients supported their theory that brain-damaged individuals are characterized by decreased cortical

conductivity. Eysenck (1955) interpreted the fact that extraverted individuals exhibited larger kinaesthetic after-effects than introverts as consistent with his hypothesis that extraverts develop reactive inhibition more rapidly than introverts. Finally, Spitz and Lipman (1961) contended that smaller kinaesthetic distortions in retardates supported the view that such subjects are less satiable than normals. In light of the present study, it is plausible to argue that these investigators measured group differences in the capacity to shift a frame-of-reference rather than differences in electro-chemical cortical activity.

This report is based on a Ph.D. dissertation submitted to the University of Alberta, Edmonton. The author is indebted to E. Howarth for his aid.

References

Charles, J. P., and Duncan, C. P. (1959). The distance gradient in kinesthetic figural aftereffect. *J. exp. Psychol.*, **57**, 164–70.

Eysenck, H. J. (1955). Cortical inhibition, figural aftereffect, and theory of personality. *J. abnorm. soc. Psychol.*, **51**, 94–106.

Gibson, J. J. (1933). Adaptation, after-effect, and contrast in the perception of curved lines. *J. exp. Psychol.*, **16**, 1–31.

Guilford, J. P. (1954). *Psychometric Methods.* New York: McGraw-Hill.

Helson, H. (1964). *Adaptation-Level Theory: An Experimental and Systematic Approach to Behavior.* New York: Harper & Row.

Howarth, E. (1964). An apparatus for measuring kinesthetic judgments. *Amer. J. Psychol.*, **77**, 492.

Klein G. S., and Krech, D. (1952). Cortical conductivity in the brain injured. *J. Pers.*, **21**, 118–48.

Köhler, W., and Dinnerstein, D. (1947). Figural after-effects in kinaesthesis. In Michotte, A. *Miscellanea psychologica.* Louvain: Catholic University of Louvain, 196–220.

Köhler, W., and Wallach, H. (1944). Figural aftereffects: An investigation of visual processes. *Proc. Amer. philos. Soc.*, **88**, 269–357.

Krauskopf, J., and Engen, T. (1960). The distance-paradox in kinesthetic figural aftereffects. *Amer. J. Psychol.*, **73**, 298–301.

Nachmias, J. (1953). Figural aftereffects in kinesthetic space. *Amer. J. Psychol.*, **66**, 609–12.

Spitz, H. H., and Lipman, R. S. (1961). A comparison of mental retardates and normals on kinesthetic figural aftereffects. *J. abnorm. soc. Psychol.*, **62**, 686–7.

Manuscript received 16th July, 1966.

PROACTIVE INHIBITION AS AN ARTEFACT OF THE METHOD OF PACED ANTICIPATION

BY

J. K. CLARKSON

From the Department of Psychology, University of Sheffield

Sixty-one subjects partly learned serial lists of 12 two-syllable adjectives presented for a set number of learning trials on two successive days. Recall was obtained 24 hr. later. One group of 30 subjects learned by the method of paced anticipation, the remaining 31 learned by silently watching the items. Free recall within a period of 1 min. was used both as a criterion for original learning and for subsequent retention. On three independent criteria for the existence of proactive inhibition (PI) the paced anticipation group showed significant PI and the silent group did not. It is suggested that PI was due to the identity of learning method, probably unique in a näive subject's experience, rather than similarity of verbal material. It is also hypothesized that difficulty in interpreting conflicting data from existing PI studies may be due to the fact that learning of verbal material is confounded with learning the S-R skill of paced anticipation.

It is only quite recently that the suitability of S-R theory as a framework within which to interpret experiments on verbal learning has been seriously challenged (Underwood, 1963; Jensen and Rohwer, 1965). The experimental methodology arising from S-R theory, however, with the rigorous control that it yields over the learning situation, is still universally employed. Subjects are presented with a stimulus word or nonsense syllable and required to make the appropriate verbal response both in paired-associate and serial learning. Generally the subject is paced, and the ability to correctly anticipate items is taken as the criterion for learning.

The concept of proactive inhibition (PI) has been largely developed on the basis of findings from numerous experiments employing this methodology. It is the aim of the present study to investigate how far method of learning may itself be a causative factor in producing interference from prior to subsequent learning.

Basic procedure follows a conventional pattern with minor variations. Subjects are presented with a serial list of 12 two-syllable adjectives for a set number of learning trials followed by a period of 1 min. allowed for free recall. Twenty-four hours later they are again asked to free recall as much of the list as they can remember in 1 min. After a brief rest the same number of learning trials for a further list is given and followed by 1 min. free recall. On day 3, 24 hr. later, they are asked to recall the second list. One group of subjects is asked to learn by attempting to verbally anticipate each item before it is presented. The other group is instructed to learn by watching the items without the requirement of an overt response.

The main variation in conventional procedure here is in the use of free recall immediately after the experimental session as a criterion of original learning, and subsequent free recall as a measure of retention. This variation was introduced in order to make treatment of the two groups exactly the same except for method of learning. It has the advantage of approaching more closely our everyday criteria for learning than does the ability to pace-anticipate items with the forced curtailment of response latencies that such a method necessitates. The period of 1 min. was decided upon on the basis of findings from a pilot study. Subjects are able to give more or less ordered recall of most items remembered within 20 to 30 sec. Rarely do they add more than two items in the remaining time. If asked 5 sec. before the end of the period if any more items can be remembered they almost invariably say "No."

Possibility of uncontrolled and uncontrollable rehearsal is a difficult problem in experiments such as this. It was decided to obtain the only evidence available in the circumstances and although strict instructions were given not to rehearse, subjects were asked at the end of the experiment if in fact they had rehearsed either list and if so, how much.

Method

Subjects sat approximately 2 ft. in front of a 12 in. × 8 in. translucent screen from behind which slides were projected. The surround to the screen was opaque and brightness was adjusted to a comfortable level. The projector was fitted with a simple device which changed slides automatically at regular intervals. Exposure time was 1·2 sec. and the screen was blank during a 0·8-sec. period required for automatic change. Two lists of 12 two-syllable adjectives were compiled from tables given by Hilgard (1951). All 24 items were chosen to be as dissimilar as possible. Words, as they appeared to the subject, were clearly projected in capital letters 1 in. high.

Subjects

Sixty-one first year psychology students, male and female, members of Stanford University were enlisted as subjects. None of them had participated in experiments of a similar nature before.

Conditions

Thirty subjects were assigned to the paced-anticipation (PA) condition and 31 to the silent (S) condition alternately as they arrived at the laboratory with the constraint that male-female ratios were kept equal over groups. Fourteen members of each group were given five learning trials for each list and the remainder eight trials. Order of learning the two lists was reversed for half of each of the four groups thus obtained. (The odd subject was assigned to the S group with eight learning trials.)

Procedure

On the first day subjects were told how many presentations of the list were to be given and instructed in the learning procedure to be followed. They were also informed that 1 min. for free recall of as many items as they could remember would be allowed after the appropriate number of learning trials. At the end of the session they were informed that free recall of the list would be required on the following day and earnestly requested not to rehearse. Day 2 commenced with 1 min. free recall of list 1 followed by a 5 min. break during which the experimenter chatted with the subject and arranged the apparatus. List 2 was then presented for the same number of learning trials followed by 1 min. recall. Instructions about rehearsal were repeated and it was emphasized that recall of list 2 only would be required on the following day. On day 3 the recall session only was given followed by informal questioning about rehearsal. Intertrial interval was kept constant at 20 sec. and a 20-sec. break was introduced before recall. Order of items remained the same throughout, although recall instructions did not specify that items should be recalled in order of original presentation. Recall was timed with a stopwatch. After 40 sec. information was given that 20 sec. remained; after 55 sec. the question "Do you think you can remember any more?" was posed.

Results

As expected from the pilot study 1 min. proved adequate time for recall. Only two subjects managed a further item in the last 5 sec. All others replied to the question "Do you think you can remember any more?" either with "No" or the equivalent of "I don't think so." In most cases ordered recall was given for the majority of items and many subjects fitted the odd items which occurred later in their correct places although not asked to do so.

An index of PI was calculated for each individual subject by use of the formula $(R_1/OL_1 - R_2/OL_2)$. Positive scores have been taken to indicate PI in Table II.

Differences between proportions showing PI under PA and S conditions are significant at $p < 0{\cdot}005$ ($\chi^2 = 10{\cdot}227$).

TABLE I

MEAN ORIGINAL LEARNING AND RECALL SCORES FOR SEPARATE GROUPS

Group	*Cond. of learning*	*No. of subjects*	*Order of lists*	*No. of trials*	OL_1	R_1	OL_2	R_2
1	PA	8	AB	8	10·0	8·4	11·0	4·9
2	S	9	AB	8	10·8	8·3	11·4	6·8
3	PA	8	BA	8	10·9	8·1	10·3	4·3
4	S	8	BA	8	11·3	8·5	11·3	7·1
5	PA	7	AB	5	9·6	7·4	9·7	4·9
6	S	7	AB	5	10·0	6·7	10·7	6·3
7	PA	7	BA	5	10·4	8·0	10·4	5·6
8	S	7	BA	5	10·7	7·1	10·3	5·1

TABLE II

NUMBER OF SUBJECTS UNDER PACED-ANTICIPATION AND SILENT CONDITIONS SHOWING PI

Group	*No. of subjects in group*	*No. of subjects showing + PI score*
PA (8 trials)	16	15
S (8 trials)	17	10
PA (5 trials)	14	13
S (5 trials)	14	8

TABLE III

INTERGROUP COMPARISONS OF PROPORTIONATE RECALL OF THE SAME LIST WHEN LEARNED ON DAY 1 AND DAY 2

Condition	*No. of subjects*	*No. of trials*	*Recall comparison*	*Value of U*	*Value of p*
PA	16	8	A_1 v A_2	4	$p < 0{\cdot}01$
			B_1 v B_2	9	$p < 0{\cdot}01$
S	17	8	A_1 v A_2	19	N.S.
			B_1 v B_2	23	N.S.
PA	14	5	A_1 v A_2	8	$p < 0{\cdot}02$
			B_1 v B_2	5	$p < 0{\cdot}01$
S	14	5	A_1 v A_2	17·5	N.S.
			B_1 v B_2	19·5	N.S.

A further measure of PI, independent of the above, was obtained from intergroup comparison of R/OL scores for the same list learned under the same conditions on day 1 and 2. The Mann-Whitney U-test has been used for convenience.

The number of intrusions occurring from list 1 into recall of list 2 together with number of subjects involved is given in Table IV.

Difference in proportion of subjects with recall intrusions under PA and S condition is significant; $p < 0{\cdot}02$ ($\chi^2 = 6{\cdot}28$).

Rehearsal

Thirty-eight of the 61 subjects admitted to some rehearsal after being told that rehearsal in such experimental situations is very difficult to avoid (as indeed it is).

Several people mentioned that before retiring for the night they would think of their plans for the following day, remember they had an experimental session to attend and immediately run through the list, hardly by conscious decision. As might be expected they were randomly distributed over experimental groups but the evidence suggests that there tends to be more rehearsal of the first list, when the experimental situation is novel to the subject, than of the second list.

TABLE IV

INTRUSIONS FROM LIST 1 INTO RECALL OF LIST 2

Condition	*No. of subjects*	*No. of subjects making intrusions*	*No. of intrusions*
PA	30	11	18
S	31	3	3

TABLE V

SUBJECTS' REPORTS ON REHEARSAL

Rehearsal of 1st list only	8
" 1st > 2nd	17
" 1st = 2nd	10
" 1st < 2nd	3

It can be seen from Table I that OL tends to be slightly higher for S groups and R_1 slightly lower except for groups 3 and 4. Whilst these small differences fall far short of statistical significance, either with groups taken singly or in combination, they may contribute in some measure to the marked differences between PA and S conditions shown in Tables II and III. The probability that $R_1/OL_1 > R_2/OL_2$ must increase to some extent as R_1/OL_1 approaches unity, as the freedom of R_2/OL_2 to differ randomly in value is curtailed in the positive direction. It was for this reason that a five-trial group was introduced although this apparently drastic reduction in practice still yields R/OL ratios well above the 50 per cent. level.

A few instances of reminiscence occurred in recall of both lists. They were randomly distributed over conditions.

DISCUSSION

The PI shown by those subjects who learned by the paced anticipation method is considerably more marked than is generally the case in studies of this kind. This may be due to the use of free recall as a criterion both for OL and for subsequent retention or possibly to the shorter than usual presentation time. Had the present experiment been conducted with the PA group alone there would have been no hesitation in attributing the interference demonstrated to some function of similarity between the lists of adjectives used. On three separate and relatively independent criteria, however, paced anticipation results in significant PI whereas silent watching of the items does not. What slight tendency towards PI may be exhibited by the S group can be safely attributed to possible differential rehearsal and to the elementary logical point that R_2/OL_2 has more room to differ negatively than positively when R_1/OL_1 is high.

When using the conventional S-R method of paced-anticipation it may easily be overlooked just how novel the situation is for the experimental subject. It is doubtful whether a näive subject has experienced any comparable method of learning prior to entry into the laboratory situation. He is confronted with the task not only of learning a list of items but also of learning each item to the criterion of a given short response latency.

The results of the present experiment seem to indicate that PI in this situation is almost entirely, if not solely, a function of learning method. It is not that list 1 interferes with list 2 through similarity of verbal content, but through identity of a unique learning experience common to both.

It has been argued in criticism of the above interpretation that since no control was exercised over order of recall, anticipation may have led to traditional serial learning while silent study may not. Hence PI for the anticipation group could be due to serial position identity. However, Underwood (1949) in his explanation of why Waters (1942) *failed* to find PI says ". . . Specific interfering tendencies must have arisen largely from serial position identity and this is known now to be an unfavourable situation for the production of PI"

Since the present experiment was conducted a comparable result for positive transfer has been demonstrated (Postman and Schwartz, 1964). The difference in method investigated was between serial and paired-associate learning. The authors conclude "The method of practice used in the first task consistently influenced subsequent learning more than did the class of verbal materials." Difficulty in interpreting and reconciling the bewildering multiplicity of existing PI studies (Slamecka and Ceraso, 1960) may, it is suggested, be partly due to the confounding of two distinct learning tasks within many such studies. The present experiment suggests the view that learning a list of words or nonsense syllables is a skill separable from that of acquiring the ability to pace anticipate.

This research was conducted at Stanford University, California, while the author was working as research associate with Dr. J. A. Deutsch, and briefly mentioned in a paper read to the Northern Branch of the British Psychological Society in October, 1962. An abstract was subsequently published (*Brit. psychol. Soc. Bull.*, Vol. 16, 1963). The author gratefully acknowledges his indebtedness to Dr. Deutsch for many stimulating discussions on the concept of PI.

References

Clarkson, J. K. (1963). Is proactive inhibition a myth? *Brit. psychol. Soc. Bull.*, **50**, 42–3.

Hilgard, E. R. (1951). Methods and procedures in the study of learning. In Stevens, S. S. (Ed.) *Handbook of Experimental Psychology*. New York: Wiley. Pp. 548–51.

Jensen, A. R., and Rohwer, W. D. (1965). What is learned in serial learning? *J. verb. Learn. verb. Behav.*, **4**, 62–72.

Postman, L., and Schwartz, M. (1964). Studies of learning to learn. 1. Transfer as a function of method of practice and class of verbal materials. *J. verb. Learn. verb. Behav.*, **3**, 37–49.

Slamecka, N. J., and Ceraso, J. (1960). Retroactive and proactive inhibition of verbal learning. *Psychol. Bull.*, **57**, 449–75.

Underwood, B. J. (1963). Stimulus selection in verbal learning. In Cofer, C. N., and Musgrave, B. S. (Ed.) *Verbal Behaviour and Learning. Problems and Processes.* New York: McGraw-Hill, Pp. 33–48.

Underwood, B. J. (1949). Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol.*, **39**, 24–34.

Waters, R. H. (1942). Degree of learning and proactive inhibition in retention. *Psychol. Bull.*, **39**, 495–6.

Manuscript received 17th August, 1966.

SHORTER ARTICLES AND NOTES

AN OPTIMUM INTERVAL IN THE ASSESSMENT OF PAIN THRESHOLD

BY

DIANA R. HASLAM and E. A. C. THOMAS

From the Department of Psychology, University of Bristol, and Statistical Laboratory, University of Cambridge

When the Limiting Method is used to measure heat-pain threshold, the observed threshold has been shown to depend on the size of the stimulus increment used (Haslam, 1965). An experiment is reported here which repeats a finding of the experiment referred to above that the variability of threshold is relatively large when the stimulus increment is small. A statistical analysis of the data shows that the previously untested hypothesis that pain threshold is uniform over the population (Hardy, Wolff and Goodell, 1952) is a reasonable one. The psychological implications of threshold variability are discussed in the light of a theoretical model, and a criterion for an optimum interval in the assessment of heat-pain threshold is discussed.

Introduction

The choice of a psychophysical procedure for the measurement of heat-pain threshold is restricted by the danger of injuring the subject if the intensity of the stimulus is too great. The procedure almost invariably adopted is the Limiting Method, in which the experimenter increases the variable stimulus intensity by discrete steps from a value well below threshold and stops when the subject first reports a pricking pain at the termination of a stimulus. A statistical analysis of this method has shown that the stopping-point is highly sensitive to the step-size used and, therefore, a correction should be made before the results of different threshold experiments are comparable (Brown and Cane, 1959). This point seems to have received little attention by experimenters until recently, when it was illustrated by Haslam (1965) who, in addition, showed that the variability of threshold is relatively large when a small step-size is used.

The basic assumptions of the Brown and Cane approach are, firstly, that the subject's sensation of a stimulus is a random variable with some distribution, which we shall call the noise distribution, and, secondly, that the subject reports pain if the sensation exceeds a certain threshold. When specific assumptions are made about the noise distribution it is possible to obtain an approximate correction formula for the observed threshold, which formula depends on the step-size and the experimental variance of a single subject. With the data available it has not been possible to use the single subject variance since most subjects were used only once. However, Thomas (1966), assuming that the noise distribution is the same throughout certain groups of subjects, has used the group variance in the correction formula to adjust the mean group threshold. It was then argued that the fact that the corrected group thresholds were more consistent than the uncorrected ones might be interpreted as evidence supporting the hypothesis of homogeneous groups.

In studying the effect of step-size on observed threshold Haslam (1965) used three step-sizes, 8, 16 and 32 mc./sec./cm.², and the three procedures were used with three groups of subjects. It was felt, therefore, that the large variance of the "8" group may have been due to heterogeneity within it. In order to check against this possibility and also to gain some information about the single subject variance, the following experiment was carried out in which a group of subjects was tested using each of the above three step-sizes twice, thus making a total of six assessments for each subject.

Method

Subjects

Eighteen subjects (10 male and 8 female) were drawn from the academic and technical staffs and the student body of the University of Bristol. Their mean age was 22·7 years.

Apparatus and procedure

Pain threshold was assessed by means of a Hardy-Wolff (1952) radiant-heat apparatus. Details of this and the procedure used are given elsewhere (Haslam, 1965).

Three different step-sizes were used: 8, 16 and 32 mc./sec./cm.² approximately. A random design, to eliminate bias, was used to assess the pain threshold of each subject six times, twice with each of the three step-sizes, and at least two days were allowed to elapse between each assessment. Although this design leaves much to be desired in that daily fluctuations in threshold are ignored, it was chosen because a preliminary investigation suggested that thresholds are not independent if assessed in fairly rapid succession.

Results

As pointed out above, because of the experimental design it seems reasonable to regard the threshold determinations as being independent of each other. An examination of the pairs of readings for a particular step-size shows no evidence of a "learning" effect, in the sense that as many subjects have recorded higher values the second time as have recorded lower ones. Thus, in the analysis of the results the order in which the readings were taken has been ignored, and the readings at each step-size have been grouped together.

A theory for the analysis of these results has already been outlined by Thomas (1966). We assume that when a stimulus of intensity t is presented, there is a probability, $p(t)$, that pain is experienced, and we assume that

$$\begin{aligned} p(t) &= 0 && \text{for } t < t_0 \\ &= (t - t_0)/2K && t_0 \leqslant t \leqslant 2K + t_0 \\ &= 1 && t_0 + 2K < t, \end{aligned} \tag{1}$$

where t_0 and K may vary among subjects. The threshold is then given by

$$T^* = t_0 + K$$

Let us denote by T_{ijr} the rth $(r = 1,2)$ assessment of threshold of the ith subject, using a step-size s_j. Let n be the integral part of $2K_i/s_j$, and let $P_m = (n-1)(n-2)\ldots(n-m)/n^m$, $m = 0,1,\ldots, n-1$. Then the probability of stopping after exactly m steps is $(P_{m-1} - P_m)$ and it can be shown that

$$E(T_{ijr} - t_{0i}) = \frac{2K_i}{n}\sum_{m=0}^{n-1} P_m \tag{2}$$

$$\doteq K_i\sqrt{2\pi/n} \text{ for large } n, \tag{3}$$

and

$$var(T_{ijr}) \doteq 4K_i^2(2 - \pi/2)/n \equiv \sigma_{ij}^2.$$

In the usual notation of the additive model, we may write

$$T_{ijr} = \mu + \alpha_i + \beta_{ij} + \epsilon_{ijr},$$

where $\alpha_i = t_{0i}$, $\beta_{ij} = K_i\sqrt{2\pi/n}$, $E(\epsilon_{ijr}) = 0$ and $var(\epsilon_{ijr}) = \sigma_{ij}^2$.

If we now assume that σ_{ij} is independent of i, i.e. that K_i is constant, then we could test the previously untested hypothesis of Hardy, Wolff and Goodell (1952) that threshold is constant over the group by doing an analysis of variance of the results for a fixed step-size. We would effectively be testing whether $t_{0i} \equiv \alpha_i$ is constant and the equation for this situation would be

$$T_{ir} = \mu' + \alpha_i + \epsilon_{ir}.$$

Unfortunately, it is felt that such a test would be over-sensitive for the following reason. When step-sizes 16 and 32 mc./sec./cm.² were used, the majority of subjects recorded the same threshold on both occasions. This indicates that the single subject variance is small for most subjects (so that the assumption of uniform K is probably not unreasonable), but because the measurement of threshold is essentially discrete this variance is observed to be zero. Thus, the observed within subject variance probably underestimates the true value, and this would overestimate the variance ratio. A homogeneity of variance test based on $U_{ij} = (T_{ij1} - T_{ij2})^2/2$ (or on $log U_{ij}$, which would have a stabilised variance), which would test whether the K_is are equal, is also impracticable with the present amount of data since many of the U_{ij}s are zero.

If t_0 and K are the same for all subjects, then the total frequency of readings at a given intensity should be approximately equal to that predicted by the 'single subject' theory sketched above. In Table I is shown the observed and expected frequencies for four ranges at step-sizes 16 and 32 mc./sec./cm.^{2} (the readings for the smallest step-size were not considered for a reason to be taken up in the next section). There is no sure way of deciding how many testing-points lie between t_0 and the minimum observed reading, i.e. of determining the location of the distribution. This has been done by aligning the modal observed and expected frequencies. The non-significant values of χ^2 (with 2 degrees of freedom) obtained suggest that the hypothesis of uniformity of pain threshold is a reasonable one.

TABLE I

EXPECTED AND OBSERVED FREQUENCIES

Step-size in mc./sec./cm^2.	*Range in mc./sec./cm.^{2}*	*≤212*	*228*	*244*	*≥260*	χ^2
16	Expected	17·9	7·3	5·4	5·4	
	Observed	15	9	6	6	1·00 N.S.
	Range in mc./sec./cm.^{2}	*196*	*228*	*260*	*≥292*	
32	Expected	7·2	11·5	10·4	6·9	
	Observed	3	14	13	6	3·79 N.S.

TABLE II

OBSERVED AND CORRECTED MEAN PAIN THRESHOLD (ENTRIES IN MC./SEC./CM.^{2})

Step-size	*Observed mean pain threshold*	*Group standard deviation*	*Within subject standard deviation*	*Corrected mean pain threshold T^**
8	201	40·0	23·6	237
16	234	33·0	10·9	222
32	252	33·8	20·0	228

Notwithstanding the above remarks, in correcting the observed threshold to obtain the "true" threshold, T^* (see Table II), we have assumed that the noise distribution is the same throughout the group and have used the formula (Thomas, 1966)

$$T^* = T_j + \frac{1{\cdot}16}{s_j} var(T_j) - 1{\cdot}91\sqrt{var\ (T_j)} \tag{4}$$

TABLE III

	Mean		*Variance*	
n	*Exact*	*Approx*	*Exact*	*Approx*
5	0·50	0·56	0·05	0·09
10	0·37	0·40	0·03	0·04
20	0·26	0·28	0·02	0·02

where $\bar{T}_j$ is the group mean and $var(T_j)$ is the within subject variance using step-size s_j, and not the group variance, as previously assumed.

The derivation of this formula depends on the approximations in equations (2) and (3). An inspection of the data shows that a reasonable value of K for the group would be 80 mc./sec./cm.², and in Table III we show the exact and approximate values of the mean and variance of $(T_{ijr} - t_o)/2K$ for $n = 5$, 10 and 20, corresponding to $s_j = 32$, 16 and 8 mc./sec./cm.². We see that the approximations are fairly good except that for $n=5$ the approximate variance is about twice the exact value. However, replacing $var(T_j)$ by $2var(T_j)$ in equation (4) when $s_j = 32$ reduces T^* by only 1 mc./sec./cm.², so that the discrepancies cancel out and the formula seems to work for small n.

From equation (3) we would expect the within subject variance to increase with the step-size. This prediction is not borne out by the data, which shows that this variance is largest when a step-size of 8 mc./sec./cm.² is used. The significance of this finding is discussed in the next section.

Discussion

The experiment reported here demonstrates a feature suggested by an earlier experiment (Haslam, 1965), that the group variance is largest when the step-size 8 mc./sec./cm.² is used; but even more important is the finding that the relative contribution to this variance of the within subject variance is largest at this step-size. This suggests that factors not taken into account in the theory outlined above become important when the step-size is very small, and that these factors tend to increase both the between and within subject variances. Within the range of validity of the theory one would expect the within subject variance to increase with the step-size, hence these results suggest that the theory ceases to be valid when the interval is very small. Thus, there seems to be an interval below which the theory is invalid and the variability of threshold is greater, and above which the theory can be applied and would predict increased variability. Since minimum variance is a desirable requirement for any experimental procedure, we shall call this interval the optimum interval for the assessment of pain threshold. The data suggest that this optimum interval lies between 16 and 20 mc./sec./cm.².

It is felt that the factors which are operative when very small intervals are used are psychological, as opposed to statistical, in nature. If the experimenter uses the same starting point for all step-sizes, then the number of trials before a near-threshold stimulus is presented increases as the step-size decreases. It may well be that some subjects are prepared to tolerate only a limited number (the "tolerance limit") of trials, so that when the interval is small, this impatience might cause the subject to end the experiment by reporting pain at a decidedly sub-threshold value. The assumption that the tolerance limit is a random variable could explain the increased within subject variance, and an assumption about the varying degrees of impatience among subjects could explain the increased between subject variance. If this interpretation is correct, one would expect the impatient subjects to record lower thresholds than the patient ones. It is perhaps relevant to consider the finding (Haslam, 1965) that, whereas introverts were found to have a lower mean pain threshold than extraverts when tested with intervals of 16 and 32 mc./sec./cm.², the reverse was the case when an interval of 8 mc./sec.²/cm. was used. A possible explanation of this is that extraverts tend to be more impatient than introverts, and it is this characteristic which is important when small stimulus intervals are used.

An alternative explanation may be as follows. With each successive presentation of the stimulus the subject is likely to believe that there is an increasing probability that he will experience pain. At the same time the suggestion of pain implicit in the instructions becomes progressively stronger. If a subject is particularly suggestible, therefore, he may report pain after relatively few stimulus presentations, and this will result in a low threshold value, especially when the step-size is small. As argued above with regard to the tolerance limit, the assumption that suggestibility is a random variable could explain the increased within subject variance, and it is well known that there are very large individual differences in degree of suggestibility, and this would explain the increased between subject variance. If suggestion is indeed an important factor when small intervals are used, then the introversion/extraversion results mentioned above could be explained by the generally accepted observation that extraverts are more suggestible than introverts.

We have assumed, with some statistical justification, that the pain threshold is uniform over the population. More evidence in support of this assumption may be derived from the fact that the estimated T^* for intervals 16 and 32 mc./sec./cm.² are 222

and 228 mc./sec./$cm.^2$ respectively, which are consistent with the value 220 mc/sec./$cm.^2$ previously suggested (Thomas, 1966).

We are greatly indebted to Mr. G. M. Clarke, now of the University of Sussex and formerly of the University of Bristol, for advice, help and discussion, and to the referee for suggesting improvements to an earlier draft of the paper. One of the authors (E.A.C.T.) wishes to acknowledge joint financial support from the Guyana Government and the University of the West Indies.

References

Brown, J., and Cane, V. R. (1959). An analysis of the limiting method. *Brit. J. statist. Psychol.*, **12**, 119–26.

Hardy, J. D., Wolff, H. G., and Goodell, H. (1952). *Pain Sensations and Reactions.* Baltimore: Williams & Wilkins Co.

Haslam, D. R. (1965). The influence of stimulus scale-interval upon the assessment of pain threshold. *Quart. J. exp. Psychol.*, **17**, 65–8.

Thomas, E. A. C. (1966). On determining pain thresholds using the limiting method. *Quart. J. exp. Psychol.*, **18**, 270–2.

Manuscript received 23rd July, 1966.

IDENTIFICATION OF CONSONANTS AND VOWELS PRESENTED TO LEFT AND RIGHT EARS*

BY

DONALD SHANKWEILER and MICHAEL STUDDERT-KENNEDY†

From the Haskins Laboratories, New York, N.Y.

The results of earlier studies by several authors suggest that speech and nonspeech auditory patterns are processed primarily in different places in the brain and perhaps by different modes. The question arises in studies of speech perception whether all phonetic elements or all features of phonetic elements are processed in the same way. The technique of dichotic presentation was used to examine this question.

The present study compared identifications of dichotically presented pairs of synthetic CV syllables and pairs of steady-state vowels. The results show a significant right-ear advantage for CV syllables but not for steady-state vowels. Evidence for analysis by feature in the perception of consonants is discussed.

Introduction

Several lines of evidence suggest that speech perception is characterized by a process different from that for the perception of other sounds (see, for example, House *et al.*, 1962; Kozhevnikov and Chistovich, 1965; Liberman *et al.*, 1966). Recent work indicating that speech and nonspeech are processed primarily in different places in the brain strengthens this hypothesis. Studies by Kimura (1961*a*) and others (Bryden, 1963; Broadbent and Gregory, 1964) have shown that if random sequences of digits are dichotically presented (Broadbent, 1954) so that a different digit arrives at each ear at the same time, listeners retain digits presented to the right ear more accurately than those presented to the left. If the stimuli are brief melodies (Kimura, 1964) or sonar signals (Chaney and Webster, 1965), listeners retain more accurately those presented to the left ear. These lateral differences in efficiency of handling competing stimuli seem to reflect the greater strength of the crossed auditory pathways (Rosenzweig, 1951) and the specialization of the auditory areas of each hemisphere of the brain for processing different classes of stimuli. Studies of persons with left- and right-sided lesions of the auditory cortex have provided independent evidence of functional differences between the hemispheres in auditory perception (Kimura, 1961*a*; 1961*b*; Milner, 1962; Shankweiler, 1966).

The relative contribution of each cerebral hemisphere to the perception of different classes of sounds merits further study. A basic question is at what stage, in the processing of speech, hemispheric differences in function become evident. Do they appear at the level of phonetic structure as well as at higher levels? Data of Chaney and Webster (1965) suggest that they do. We may then ask whether all phonetic elements, or all features of the phonetic elements, are processed in the same way. The technique of dichotic presentation offers an approach to these questions in persons with intact nervous systems, and one by which we may hope to learn more explicitly how speech perception differs from the perception of other sounds.

In the present study we compared identification of dichotically presented single pairs of synthetic steady-state vowels and of consonant-vowel (CV) syllables. We also examined the effects of interaural competition for the various combinations of articulatory features contained in the consonant syllable pairs.

Method

Two separate tests were made up, one consisting of synthetic consonant-vowel syllables and the other of synthetic steady-state vowels in isolation.

Six consonant-vowel syllables lasting 300 millisec., were prepared on the Haskins Pattern Playback (Cooper *et al.*, 1951) and recorded on magnetic tape. (The Pattern Playback is a device for producing controlled synthetic speech stimuli; hand-painted

* An early version of this work was presented at the 71st meeting of the Acoustical Society of America, Boston 1st–4th June, 1966 (*J. acoust. Soc. Amer.*, **39**, 1256, Abstract).

† Now at Inter American University, San German, Puerto Rico.

patterns, modelled on spectrograms, are converted photo-electrically into sound.) They were the voiced stop consonants [b, d, g] and the unvoiced stop consonants [p, t, k], each followed by the vowel [a]. The fifteen possible syllable pairs (no syllable was paired with itself) were made into loops, aligned for simultaneity of onset and re-recorded on a dual channel tape recorder (Ampex PR–10). Each member of a pair was recorded twice on channel A and twice on channel B and the resulting 60 stimulus pairs were spliced into a random order.

A similar procedure was followed with five, equal duration (300 millisec.), steadystate vowels [i, ϵ, æ, a, u]. The 10 possible pairs, with each member of a pair recorded twice on each channel yielded a random order of 40 pairs.

The subjects for the consonant test were 15 right-handed members of the laboratory staff, 10 women, five men between the ages of 20 and 50. Ten of these listeners also served as subjects in the vowel test. None had a known hearing loss.

Subjects were tested individually in a quiet room. They listened over earphones (Permoflux, PDR-8), wired for dichotic presentation, to the output from the tape recorder. The subjects operated the recorder themselves, starting and stopping the tape before and after each trial, and listening to each trial pair only once. They wrote their judgements on an answer sheet at the top of which were displayed the six consonants (or five vowels) from which they were to make their choices. Listeners were instructed that each stimulus pair would contain two different syllables and that they were to attempt to identify both, guessing if necessary. They were asked to write the judgement of which they were more confident in the first column and their other judgement in the second column. Each listener heard each test twice, the earphones being reversed on the second run, so that channels and ears were balanced. Approximately half the listeners heard channel A in their left ear first; half heard channel B in their left ear first.

Results and Discussion

Consonant-vowel comparison

The results of the consonant test showed that syllables presented to the right ear were identified with greater accuracy than were those presented to the left by 14 of the 15 subjects. The figures given below are mean per cent. correct for each ear when only the response given first is counted.

Left ear	Right ear
29 per cent.	45 per cent.

The difference is highly significant ($p < 0{\cdot}001$ by a two-tailed test). First preferences only are recorded since it turned out that second preferences were largely guess work.

TABLE I

Mean per cent. Correct on First Preferences: Comparison of Consonants and Vowels: Ten Subjects

	Left ear	*Right ear*	*P*
Vowels	41	45	Not significant
Consonants ..	31	45	$<0{\cdot}001$

The right ear advantage shows up again when we compare the results of the consonant and vowel tests for the 10 subjects who took both. Table I shows first preference figures for these two conditions. Right ear performance is identical for both consonants and vowels—45 per cent. correct. Left ear performance is slightly but not significantly, lower for the vowels (41 per cent.), but is significantly lower for the consonants (31 per cent.) ($p < 0{\cdot}001$ by a two-tailed test). The right ear advantage occurred unreliably on the vowel test (six of the 10 subjects were better on the right ear).

In view of Kimura's finding (1964) of a left ear advantage for musical melody recognition, as against a right ear advantage for spoken digits, the neutral status of steady state vowels, midway, as it were, between speech and music, is perhaps not surprising. Whether this status will be maintained by vowels placed in dynamic context remains to be seen.

The role of articulatory features in identification of consonants

The set of six stop consonant syllables can be classified into groups according to features of voicing (which has two values: present, absent) and place of articulation (which has three values: labial, alveolar, velar). This breakdown is shown in Table II.

TABLE II

PAIRED COMBINATIONS OF SIX STOP CONSONANTS ACCORDING TO FEATURES OF VOICING AND PLACE OF ARTICULATION

	Place of articulation		
	Labial	*Alveolar*	*Velar*
Unvoiced	p	t	k
Voiced	b	d	g

Pairs differing in

Voicing and place	*Voicing alone*	*Place alone*
p—d p—g t—b t—g k—b k—d	p—b t—d k—g	p—t p—k t—k b—d b—g d—g

Of the 15 possible syllable pairs, six contrast in two features (voicing and place), nine contrast in one feature (three in voicing, but not in place; six in place, but not in voicing).

The percentages of correct responses on first preferences for these feature contrasts are given for each ear in Table III. Right ear performance is the same for all three

TABLE III

SYNTHETIC CV SYLLABLES: FIRST PREFERENCE MEAN PER CENT. CORRECT FOR EACH EAR ACCORDING TO FEATURE DIFFERENCES: FIFTEEN SUBJECTS

Pairs differing in:					
Voicing and place		*Voicing alone*		*Place alone*	
Left	*Right*	*Left*	*Right*	*Left*	*Right*
11	28	17	29	22	28

feature contrast conditions; it is only for the left ear that the conditions differ in difficulty. The resulting ear differences show a marked trend. The advantage is greatest for the double contrast (17 per cent.), next greatest for the voicing contrast (12 per cent.), least for the place contrast (6 per cent.). The overall effect is significant by analysis of variance with $p < 0{\cdot}025$. The difference in right ear advantage for the contrast conditions of voicing alone and place alone is not significant, but the difference between the average of these two and the double contrast is significant with $p < 0{\cdot}05$. There is nothing in auditory psychophysics that would lead one to predict that simultaneous stimuli differing along two feature dimensions should be more difficult to identify than stimuli differing on

one. This outcome supports other studies, such as those reported by Kozhevnikov and Chistovich (1965), suggesting that the perception of consonant-vowel syllables may involve some process of analysis by feature.

The effect is due, as we saw, to the greater error rate for syllables presented to the left ear when two features vary, suggesting that speech signals of this kind are more readily processed in the left hemisphere when the identification task is difficult. A detailed analysis of the errors is needed to discover precisely what the left hemisphere does better than the right. This will require a confusion matrix analysis which we are currently carrying out on a larger body of data collected on dichotically presented natural speech syllables. The analysis may throw some light on the feature system used in the perception of consonants. The system may then be compared with that used in memory as exemplified in recent experiments by Conrad (1964) and Wickelgren (1966). A masking experiment such as that of Miller and Nicely (1955) does not speak to this issue because the masking stimulus has an unequal effect on different features.

Conclusions

We can summarize the main findings and implications of this experiment as follows: (1) Relatively large and stable laterality effects occur on dichotic presentation of nonsense syllables displaying phonemic contrasts. This strongly suggests that left hemisphere dominance in speech perception operates at the level of speech sound structure. (2) The effect can be demonstrated when only a single pair of syllables is presented on each trial, indicating that it pertains to the registration of the stimuli and not only to their retention. (3) The effect is significant for synthetic consonant syllables, but not for synthetic steady-state vowels. (4) The effect is greater for consonant-vowel pairs differing on two articulatory features than for pairs differing on one. This suggests that the perception of such consonant syllables may involve a process of analysis by feature.

This work was supported in part by a grant to Haskins Laboratories from the National Institute of Child Health and Human Development.

References

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *J. exp. Psychol.*, **47**, 191–6.

Broadbent, D. E., and Gregory, M. (1964). Accuracy of recognition for speech presented to the right and left ears. *Quart. J. exp. Psychol.*, **16**, 359–60.

Bryden, M. P. (1963). Ear preference in auditory perception. *J. exp. Psychol.*, **65**, 103–5.

Chaney, R. B. and Webster, J. C. (1965). Information in certain multidimensional acoustic signals. *U.S. Navy Electronics Laboratory Reports* (San Diego, Calif.), No. 1339.

Conrad, R. (1964). Acoustic confusions in immediate memory. *Brit. J. Psychol.*, **55**, 75–84.

Cooper, F. S., Liberman, A. M., and Borst, J. M. (1951). The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proc. Nat. Acad. Sci.*, **37**, 318–28.

House, A. S., Stevens, K. N., Sandel, T. T., and Arnold, J. B. (1962). On the learning of speechlike vocabularies. *J. verb. Learn. verb. Behav.*, **1**, 133–43.

Kimura, D. (1961*a*). Cerebral dominance and the perception of verbal stimuli. *Canad. J. Psychol.*, **15**, 166–71.

Kimura, D. (1961*b*). Some effects of temporal-lobe damage on auditory perception. *Canad. J. Psychol.*, **15**, 156–65.

Kimura, D. (1964). Left-right differences in the perception of melodies. *Quart. J. exp. Psychol.*, **16**, 355–8.

Kozhevnikov, V. A., and Chistovich, L. A. (1965). *Rech' Artikuliatsia i Vospriiatie.* Moscow-Leningrad. Tr. *Speech: Articulation and Perception.* Joint Publications Research Service, Washington, D.C., JPRS: 30,543.

Liberman, A. M., Cooper, F. S., Studdert-Kennedy, M., Harris, K. S., and Shankweiler, D. P. (1966). Some observations on the efficiency of speech sounds. *Proc. XVIII Intern. Congr. Psychol.*, Moscow, in press.

Miller, G. A., and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. acoust. Soc. Amer.*, **27**, 338–52.

MILNER, B. (1962). Laterality effects in audition. MOUNTCASTLE, V. B. (Ed.), *Interhemispheric relations and cerebral dominance.* Baltimore: Johns Hopkins Univ. Press.

ROSENZWEIG, M. R. (1951). Representations of the two ears at the auditory cortex. *Amer. J. Physiol.*, **167**, 147–58.

SHANKWEILER, D. (1966). Effects of temporal-lobe damage on perception of dichotically presented melodies. *J. comp. physiol. Psychol.*, **62**, 115–9.

WICKELGREN, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. *J. acoust. Soc. Amer.*, **39**, 388–98.

Manuscript received 22nd August, 1966.

ON PROCURING HUMAN SUBJECTS

BY

D. R. J. LAMING*

From the Department of Psychology, University College, London

A large number of subjects was required for a series of choice-reaction experiments. An ample supply of volunteer subjects was obtained without excessive trouble by sending a circular letter to every undergraduate in Trinity College, Cambridge. Several different types of circular were compared, but none of them proved any more effective than the others. A probabilistic model is developed which accounts adequately for the times taken by the undergraduates to reply. The discussion following covers the implications of this model, and a technical reason why the different letters should appear equally effective.

Introduction

The series of experiments reported by Laming (1967) required about 200 subjects, and procuring these subjects was a subsidiary problem in its own right. Previous experience had shown that a notice placed amongst a medley of undergraduate advertisements on a college noticeboard would not attract a sufficient number of volunteers. I therefore circulated the entire undergraduate population of Trinity College, Cambridge, of which I was a resident member at that time, with a request for help. The undergraduates were asked to volunteer as subjects for an experiment at some unspecified time in the future; it was understood that they would be approached again individually when their help was required. A similar scheme existed earlier in Liverpool under the aegis of Dr. A. Heron.

Since there were 720 men to be circulated, there was an opportunity to learn something about how best to do it. It was planned as a deliberate experiment and the details were as follows:

Experiment I

Method

Twelve different letters were used. They were distinguished by three factors: first the opening paragraph was either "emotional" (an appeal from a research student in great need of help), "formal" (a formal request for help), or "rational" (arguing that the recipient ought to come and help in the cause of science). Second, half of the letters contained a middle paragraph giving a few details of work done during the previous year (reported in Laming, 1962); the other letters did not include this. Third, half of the letters described the experiments as "psychological"; the others made no reference to psychology at all. All of the letters were headed "Trinity College, Cambridge" and ended with the same final paragraph, asking the recipient to return a reply card (enclosed) if he was willing to help.

TABLE I

Numbers of Replies (out of 60) Received to Each Letter Used in Experiment I

	With middle paragraph		*Without middle paragraph*	
	Reference to psychology	*No reference to psychology*	*Reference to psychology*	*No reference to psychology*
"Emotional"	21	22	24	20
Formal	19	18	15	21
Rational	16	21	22	22

Sixty copies of each letter were sent out. The undergraduates were grouped according to their year and whether they were reading Arts or Science and each group was distributed evenly amongst the 12 alternative letters. The letters were delivered on one day in October, 1960, and the accumulated replies were collected at the same time each day,

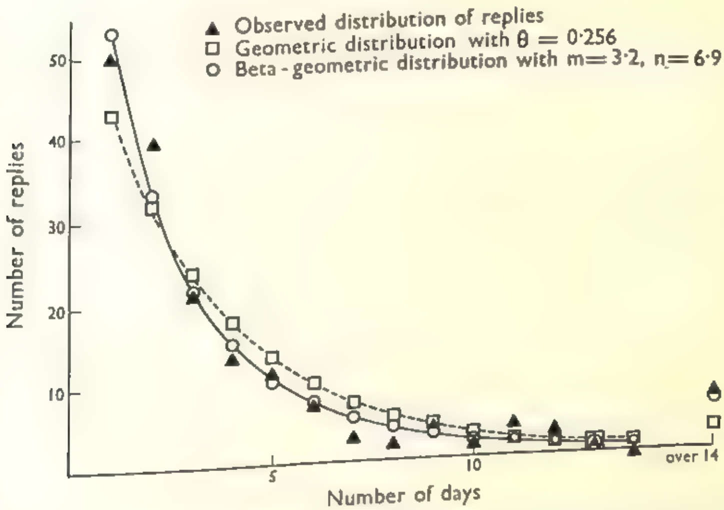
* Now at Psychological Laboratory, Cambridge.

beginning on the following day and continuing until the fourteenth day, on which day no replies were received. A few further replies came in between then and the end of term, but no date of receipt was recorded with these.

Results

A total of 241 replies was received to the 720 letters sent. The numbers of replies (out of 60) to each type of letter are shown in Table I. The differences between these numbers are far from significant, nor is there any difference which can be attributed to any of the factors distinguishing the letters. There is no difference in the proportions of replies from undergraduates of different years, but a slightly higher proportion of scientists replied than arts men ($\chi^2 = 3{\cdot}66$ with 1 *d.f.*, which is nearly significant at the 0·05 level). A greater proportion of replies was received from men living in college than from those out in lodgings ($\chi^2 = 5{\cdot}96$ with 1 *d.f.*, $p < 0{\cdot}02$), who also took longer on average to reply (5·38 days instead of 3·90 days for men in college). The temporal distribution of the replies from men living in college is shown in Figure 1, where it is compared with the best

FIGURE 1



Distribution of "In College" replies in Experiment I compared with two theoretical distributions.

fitting geometric distribution ($\hat{\theta} = 0{\cdot}256$ is the maximum likelihood estimate of the distribution parameter; it is asymptotically equivalent to a minimum χ^2 estimate). The difference between these two distributions is not quite significant ($\chi^2 = 14{\cdot}82$ with 8 *d.f.*, $p > 0{\cdot}05$).

Experiment II

In the summer of 1961 a second letter was sent to all who had replied to the first letter, thanking them for their offer of help and enquiring whether they were willing to assist in the same way during the following year. Out of 195 who were in residence the following year 125 replied. There was no difference in the proportion of replies from those who had done an experiment during the year and those whose offer had not been taken up; nor was there any difference in the proportion of replies from those in college and those out in lodgings. The temporal distribution of replies from those men living in college was approximately the same as that of the replies to the first letter. The maximum likelihood estimate of θ is 0·230 and the observed distribution is significantly different from the best-fitting geometric one ($\chi^2 = 22{\cdot}68$ with 7 *d.f.*, $p < 0{\cdot}01$). The difference consists in too large a proportion of replies on the first and subsequent to the eighth days. Finally, the Spearman rank correlation coefficient between individual reply times to the first and second letters was 0·223; this is significant at the 0·05 level ($t = 2{\cdot}01$ with 77 *d.f.*; see Siegel, 1956, pp. 202–13).

Experiment III

Method

Eighteen months later further subjects were needed. A third set of letters was sent out to the then first and second year undergraduates in Trinity College, who had not previously been circulated. Eight different letters were used. They were similar to the "formal" letters of the first set and differed amongst themselves in the following respects:

(i) by the substitution of "Dear Sir" and "Yours faithfully, D. R. J. Laming" for "Dear Mr. . . . " and "Yours sincerely" followed by a signature;

(ii) by the omission of the middle paragraph as in the first set of letters;

(iii) by the substitution of the address "The Psychological Laboratory, Downing Street, Cambridge" at the head of the letter in place of "Trinity College, Cambridge."

Fifty-five copies of each letter were sent out. The undergraduates were grouped according to their year and whether they lived in college or in lodgings and each group was distributed evenly amongst the eight alternative letters. The letters were delivered in one day in January, 1963, and the accumulated replies were collected at the same time each day until the eighteenth day.

TABLE II

Numbers of Replies (out of 55) Received to Each Letter used in Experiment III

Address at head of letter	*With middle paragraph*		*Without middle paragraph*	
	Psychological Laboratory	*Trinity College*	*Psychological Laboratory*	*Trinity College*
"Dear Sir," etc. ..	14	12	14	17
"Dear Mr.—," etc. ..	18	15	14	15

Results

One hundred and nineteen replies were received to the 440 letters sent. This proportion is slightly less than that of October, 1960 ($\chi^2 = 4{\cdot}97$ with 1 *d.f.*, $p < 0{\cdot}05$). The numbers of replies (out of 55) to each type of letter are shown in Table II. The differences between these numbers are again far from significant, nor is there any difference which could be attributed to any of the factors distinguishing the letters. There is no difference between the proportions of replies from undergraduates of different years, nor between the proportions received from those living in college and those out in lodgings. The temporal distribution of replies is shown in Figure 2. It is similar to the distributions of replies to the first two sets of letters and differs from the best fitting geometric distribution ($\hat{\theta} = 0{\cdot}239$) in the same way ($\chi^2 = 19{\cdot}43$ with 9 *d.f.*, $p < 0{\cdot}05$).

The Distribution of Reply Times

The geometric distribution is important in this context. Suppose that for each undergraduate who ultimately replies the conditional probability of sending a reply on a particular day, given that no reply has hitherto been sent, is constant, say θ. Then the number of days taken by an undergraduate to reply will be a geometric random variable with parameter θ. But the distribution of replies from the population of undergraduates as a whole will only be geometric if the parameter is constant throughout the population. Evidently this is not so.

Suppose instead that the parameter θ has a beta distribution, $Be(m,n)$ in the population of undergraduates who replied:

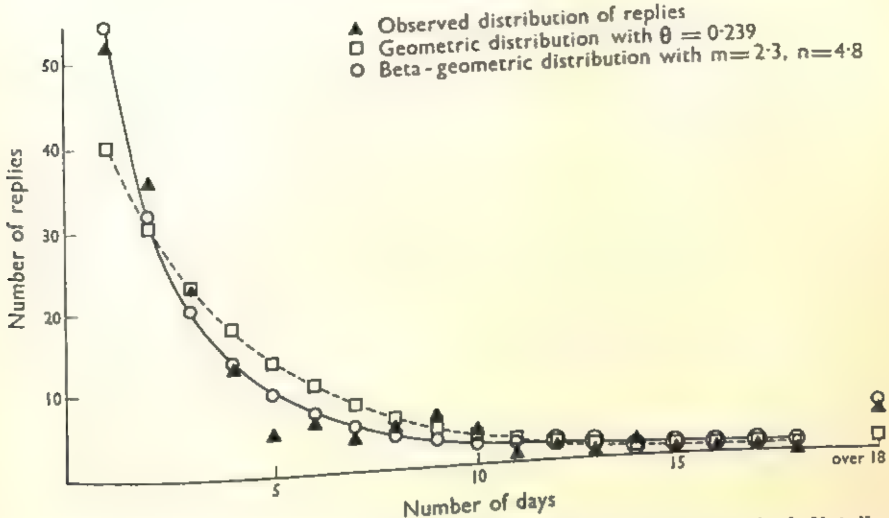
$$\text{i.e. } f(\theta) = \frac{(m+n-1)!}{(m-1)!\,(n-1)!}\,\theta^{m-1}(1-\theta)^{n-1} \tag{1}$$

Then, if g_r denotes the expected proportion of replies on the r^{th} day,

$$g_r = \frac{m.n(n+1)\,\ldots\ldots\ldots\ldots\,(n+r-2)}{(m+n)(m+n+1)\,\ldots\ldots\,(m+n+r-1)} \tag{2}$$

The beta distribution is introduced here for practical convenience rather than for any theoretical reason. Necessarily the parameter of a geometric distribution lies in the interval $0 < \theta < 1$, and since the two parameters of the beta distribution may assume any positive real values, most unimodal distributions on the unit interval may be approximated by a suitable choice of m and n. Most of the statistics of interest may be obtained easily as beta integrals.

FIGURE 2



Distribution of replies in Experiment III compared with two theoretical distributions.

In practice the number of replies arriving on each day was recorded for a certain number of days only. Let ν_r, $r = 1 \dots k$, be the number of replies received on the r^{th} day and let ν_{k+1} be the total number of replies received on or after the $(k+1)$th day. Then the maximum likelihood estimates of m and n satisfy

$$\left.\begin{aligned}\sum_{r=1}^{k}\left[\left(\sum_{i=r+1}^{k+1}\nu_i\right)m-\nu_r(n+r-1)\right]\Big/(m+n+r-1)&=0\\ \sum_{r=1}^{k}\left[\left(\sum_{i=r+1}^{k+1}\nu_i\right)m-\nu_r(n+r-1)\right]\Big/(n+r-1)&=0\end{aligned}\right\}\qquad(3)$$

and these equations may be solved by iteration. For the distributions of replies shown in Figures 1 and 2, equations (3) have approximate solutions (3·2, 6·9) and (2·3, 4·8) respectively for (m, n). The theoretical distributions with these parameters are shown in Figures 1 and 2 and do not differ significantly from the observed distribution in either case. The two sets of estimates (m, n) are reasonably consistent. The theoretical mean reply time is $(m+n-1)/(m-1)$ and this takes the values 4·14 and 4·67 for the two sets of data. The theoretical variance is $mn(m+n-1)/(m-1)^2(m-2)$, which takes the values 41·5 and 132·9 respectively. The difference between the estimates therefore reflects for the most part a difference in sample variance.

Discussion

It happens that exactly the same theoretical distribution of replies is obtained from a Polya Urn scheme (see Feller, 1957, p. 110, with $b = m$, $r = n$ and $c = 1$). In this case the same distribution of reply times with fixed parameters holds for every member of the population, but the longer a man puts off replying, the less likely he is to reply on the

following day. The data of Figures 1 and 2, of course, do not distinguish between these two interpretative possibilities.

However, the Polya Urn scheme predicts that two sets of replies from the same population should be uncorrelated, and in the comparison of Experiments I and II this has been found not to be so. The beta-geometric model, on the other hand, gives a theoretical correlation coefficient of $1/m$, even though successive replies from the same individual are independent. In fact the sample product-moment correlation coefficient between Experiments I and II is 0·078 (this is computed from the raw scores while the Spearman coefficient used above is computed from ranks), while the value predicted from Experiment I is 0·312. It therefore seems likely that while the parameter θ of an individual's geometric reply distribution varies in the population studied, its distribution is more sharply peaked than the estimates (m, n) (in this case 3·2, 6·9) imply.

As well as the speed of reply, it appears that the probability of replying at all also varies in the population studied. Suppose instead the contrary, namely, that the probability of replying, ϕ, is constant for each recipient of a letter. Then the number of replies to the different letters in Tables I and II should show a variation at least as great as the theoretical dispersion of binomial random variates. In fact these scores are more alike than would happen by chance, yielding unusually small values of χ^2 (Table I: $\chi^2 = 5{\cdot}77$ with 11 *d.f.*; Table II: $\chi^2 = 2{\cdot}30$ with 7 *d.f.*). The sum of these χ^2 statistics is significant at the 0·025 level, at the bottom tail of the distribution; this implies that the scores in the tables arise from mixed binomial trials, that is, that ϕ varies in the population studied. The variance of a mixed binomial statistic is less than that of a binomial with the same average ϕ (see Kendall, 1945, Vol. 1, p. 122) and it is possible to estimate the variance of ϕ in this way. Neglecting any difference of effect between the letters, Var ϕ is 0·105 for Table I and 0·135 for Table II. The difference is not significant.

One effect of the mixed binomial trials is that it is difficult to demonstrate a difference between the letters, even where one exists. The scores in Tables I and II were transformed by the arcsin transformation and re-examined by an analysis of variance (see Snedecor, 1956, §11.12), but the result was qualitatively the same.

It is possible to name some of the factors which influence the probability of a given undergraduate replying. Personal friends and acquaintances, of course, are more likely to reply. Scientists are a little more likely to reply than men reading arts subjects. But the experience of doing an experiment during the preceding year seems not to make any difference. The third set of letters drew a slightly smaller proportionate response than the first set, probably because they were sent out later in the academic year, when the undergraduates were more settled in their leisure activities. Pressure to reply or effort necessary to make a reply also affect the issue. In the design of the experiments above two principles were observed; no pressure was put on the recipients to respond (except possibly through the contents of the letters) and the action of sending a reply was made as easy and effortless as possible.

The variation of the probability of replying amongst the recipients of the letters effectively masked whatever difference in effect there might have been between the letters. Provided the letter is not deliberately bad, its exact nature is not very important. It might be possible to demonstrate that certain types of letter are better than others by determining how the probability of sending a reply varies in the population under study. But as a means of finding additional subjects such information would be of only marginal value.

The statistical characteristics of the reply times and frequencies seem but to confirm the expected. Careful thought in advance of the experiment would probably have suggested precisely those hypotheses that have been justified by the results. Of rather more value is the demonstration that it is possible to find so many subjects as this for psychological experiments without too great an expenditure of money or time. It may well be that the degree of success of this particular operation is peculiar to a residential university, or even to Trinity College, Cambridge. It seems likely, too, that the response would have been rather poorer if there had been any competition from other experimental psychologists. The letters sent out seemed to arouse considerable interest and goodwill, and the subjects obtained in this way co-operated well in some rather tedious experiments.

I am indebted to Miss V. R. Cane for several helpful comments, and in particular for drawing my attention to the Polya Urn scheme. This work was done at the Psychological Laboratory, University of Cambridge, while I held a D.S.I.R. studentship, and the paper was prepared with support from a grant provided by the S.R.C.

References

Feller, W. (1957). *An Introduction to Probability Theory and its Applications*, Vol. 1, 2nd ed. New York: Wiley.

Kendall, M. G. (1945). *The Advanced Theory of Statistics.* London: Griffin.

Laming, D. R. J. (1962). A statistical test of a prediction from information theory in a card-sorting situation. *Quart. J. exp. Psychol.*, **14**, 38–48.

Laming, D. R. J. (1967). *The Information Theory of Choice-reaction Times.* E.P.S. Monograph No. 3.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Snedecor, G. W. (1956). *Statistical Methods* (5th ed.). Iowa: Iowa State College Press.

Manuscript received 8th August, 1966.

REACTION TIMES FOR SIMPLE SHAPE DISCRIMINATIONS REQUIRING ONE OR BOTH VISUAL CORTICES

BY

G. S. BRINDLEY, R. H. S. CARPENTER and D. N. RUSHTON

From the Physiological and Psychological Laboratories, University of Cambridge

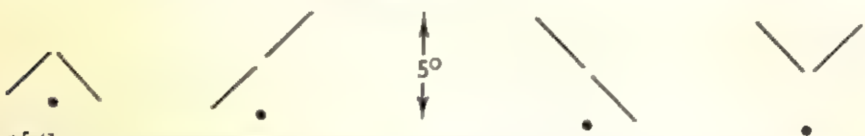
Reaction times for a simple two-choice shape discrimination requiring either one or both visual cortices were measured.

In a total reaction time of around 400 millisec. the difference found was $3{\cdot}0 \pm 2{\cdot}6$ millisec. If subjects were weighted according to number of observations, and $-1{\cdot}34 \pm 1{\cdot}68$ millisec. if they were weighted according to reciprocals of variances of differences of means; that is, it was not significant.

Introduction

It might reasonably be supposed that a shape discrimination requiring only the information received by the visual cortex of one hemisphere could in general be performed more quickly than one requiring the information received by both hemispheres. Partly

FIGURE 1



Four of the 16 stimulus cards. The small filled circle below each pair of lines is the fixation mark; thus these are two-hemisphere stimuli. The other 12 cards had patterns derived by rotating the four shown in the Figure about the fixation mark through 90° (one-hemisphere), 180° (two-hemisphere), and 270° (one hemisphere).

TABLE I

Subject	*Observations (excluding mistakes)*	*Figure all in one half-field*		*Figure split between half-fields*		*Standard error of difference of means*
		Mean (millisec.)	*Mistakes*	*Mean (millisec.)*	*Mistakes*	
1	447	394·4	11	393·2	10	3·1
2	609	423·9	24	430·8	24	3·4
3	393	549·2	6	555·8	4	6·1
4	358	510·3	9	536·1	6	8·2
5	444	506·3	10	492·0	15	7·8
6	370	498·1	30	483·8	28	10·3
7	442	429·3	29	432·4	20	9·7
8	468	475·0	10	465·2	3	9·7
9	702	470·9	20	479·1	13	7·3
10	396	411·7	4	404·0	0	13·3
11	399	420·2	5	396·5	2	13·6
12	677	396·3	34	392·5	21	8·5
13	648	424·6	14	423·6	10	9·4
14	369	473·1	3	461·4	2	17·1
15	376	348·7	15	345·1	11	16·9
16	359	356·5	8	354·4	12	17·9
17	394	393·8	4	393·6	2	16·7
18	363	449·3	5	441·1	10	18·7

on theoretical grounds and partly on the basis of preliminary experiments we judged the set of stimuli of Figure 1 to be the most likely to reveal such a difference in speed of discrimination if it existed.

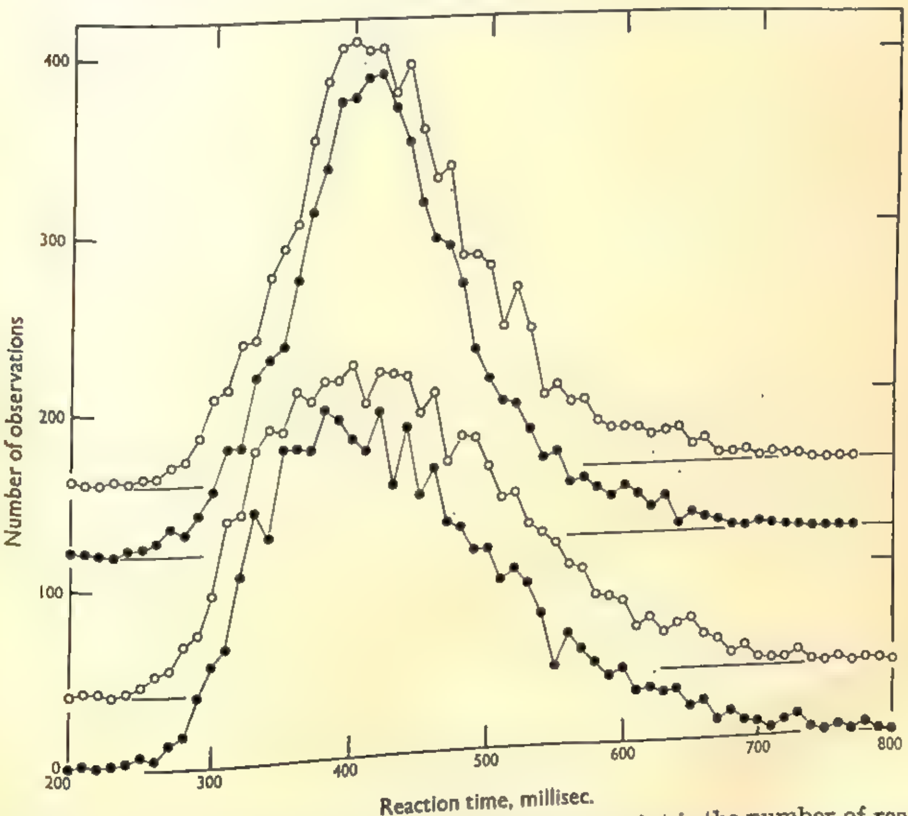
METHOD

A box tachistoscope was used, and the stimuli were drawn on cards which were exposed for 100 millisec. A fixation mark was continuously visible. The figure to which the subject had to react appeared above, below, or to one side of the fixation mark, and consisted of two short lines at 45° to the vertical forming either a straight line or a right angle, as in Figure 1. The subject viewed the stimuli monocularly through a prism, by rotation of which the entire array could be turned through 90°. By doing this between the first and second of the two similar sessions (often consecutive) for which each subject served, we controlled for any differences in drawing or illumination, since the figures that had been projected to one hemisphere only were now projected to both, and vice versa.

The subject was instructed to press one of two buttons if the two halves of the figure formed a right angle, and the other if they formed a straight line.

The set of 16 stimulus cards was presented one by one, then shuffled, then presented again, and so on. A warning click was given one sec. before the tachistoscope flash. A dekatron counter was started at the same time as the flash, and stopped by the subject's pressing one of the buttons. The very few reaction times that were less than 200 or more than 800 millisec. were discarded. Mistakes (i.e. instances where the subject pressed the wrong button) were much commoner; they were recorded and are shown in Table I.

FIGURE 2



Distributions of reaction times. The ordinate of each point is the number of reaction times that fell in the 10 millisec. interval to the right of the time shown by its abscissa. The open circles are for figures wholly within one-half of the visual field ("one-hemisphere"), the filled circles for figures divided between the two halves of the field ("two-hemisphere"). In the two lowest curves all observations are pooled without adjustment, in the upper curves with adjustment as described in the text. The upper three curves have their zeros displaced upwards by 40, 120 and 160 observations respectively.

Results

The two lowest curves in Figure 2 show the pooled distributions of reaction times for the one-hemisphere and two-hemisphere conditions. The upper curves of the same Figure show the result of computing the mean reaction time for each session of each of the 18 subjects (taking the one-hemisphere and two-hemisphere observations together), adding to every reaction time the constant needed to bring the mean for its session to the arbitrary value of 420 millisec., and then combining the 36 sessions, treating the one- and two-hemisphere observations separately. The means of the resulting combined distributions differ by 3·0 millisec., that for the two-hemisphere condition being the smaller, contrary to our expectation. The standard error of the difference of means is 2·6 millisec. (8178 *d.f.*), so the result does not differ significantly from zero.

The above is an appropriate statistical treatment if one assumes that the "true" effect of dividing the incoming information between the two hemispheres varies from subject to subject, and wishes to know the average value of this effect in the human population; all observations other than mistakes contribute equally to the final estimate, irrespective of which subject made them. However, as can easily be seen from the last column of Table I, subjects varied greatly in the scatter of their reaction times. It might be argued that a more useful estimate would be obtained by weighting the differences of means for different subjects not merely in proportion to the numbers of observations, but in proportion to the reciprocals of the variances. This would be the correct procedure if it were assumed that the "true" effect of dividing the information between the hemispheres is the same for all subjects, so that different subjects are providing independent estimates of the same quantity. The weighted mean obtained by using it is −1·34 millisec., where the minus sign indicates that the two-hemisphere condition gives the longer reaction time. The standard error of the weighted mean is 1·68 millisec., so the difference from zero is again insignificant.

Discussion

The result provides no evidence that the brain is any slower in deciding whether an interrupted line is straight or bent when the necessary information is divided between the two visual cortices than when it all comes to one; in particular, it is incompatible, for our test conditions, with a slowing greater than about 5 millisec.

Manuscript received 23rd August, 1966.

CHOICE REACTION TIME TO SINGLE DIGITS, SPELLED NUMBERS, "RIGHT" AND "WRONG" ARITHMETIC PROBLEMS AND SHORT SENTENCES

BY

JANE A. NORTH, DAVID A. GRANT and ROBERT A. FLEMING

From the Psychology Department, University of Wisconsin, Madison, Wisconsin U.S.A.

Average RT to single number stimuli were found to differ significantly, with the shorter RTs being evoked by 1, 2, 6, and 9. The relationship was most marked when the numbers were presented visually as digits but held also when the numbers were printed as words.

RT for classification of simple three-digit addition and subtraction problems as correctly or incorrectly added or subtracted was shortest for correct additions and about equal for incorrect additions and correct and incorrect subtractions, implying a difference in processing of these forms of information. Similarly, dubitably false sentences were more slowly classified as "true" or "false" than were indubitably false sentences and either dubitably or indubitably true sentences.

Introduction

The present paper reports choice reaction time of college students to single digits and spelled numerals. Also reported are the average times required to classify as "right" or "wrong" simple addition and subtraction problems and short English sentences. The data were obtained primarily to determine optimal CS-UCS temporal intervals in differential eyelid conditioning experiments, but they turned out to possess some intrinsic interest and are reported for that reason. They should be of interest to investigators following up work on RT as a function of uncertainty (e.g. Hick, 1952; Hyman, 1953; Fitts and Switzer, 1962; Fitts, Peterson, and Wolpe, 1963; Forrin and Morin, 1966); those who are studying response latencies in naming objects (e.g. Morin, Konick, Troxell, and McPherson, 1965); and investigators of information processing along the lines of Wason (1959) and Thomas (1963).

Experiment I

Method

The stimuli consisted of the digits 1 through 9, the spelled numbers ONE through NINE, and three-digit addition and subtraction problems in the form: $8 - 2 = 6$, $8 + 1 = 7$, etc. The problems might be right or wrong. If wrong, the error never exceeded 3. Letters and digits were printed in black India ink with a Leroy lettering set and were ½ in. in height.

Stimuli were presented for 2 sec. by means of a Scientific Prototype three-channel mirror tachistoscope. The viewing distance was 4 ft., and the luminance of the pre- and post-exposure fields was 70 ft. candles. Reaction times were obtained from a Hunter voice-operated relay and a Hunter Clockcounter.

The subject was instructed to look in the eye-piece of the tachistoscope and was told that digits, spelled numbers, and arithmetic problems with their answers would appear shortly after "ready" signals. He was told to state which number had appeared as quickly and accurately as possible and with the arithmetic problems he was to respond "true" or "false" depending upon whether the answer given was correct or incorrect. A subject might then see a sequence of 18 spelled numbers, a sequence of 18 digits, and a sequence of 25 arithmetic problems. The order within sequences varied randomly from subject to subject, and the order of sequences was varied across subjects. A 2 sec. fore-period was used, and the intertrial interval was 10–15 sec.

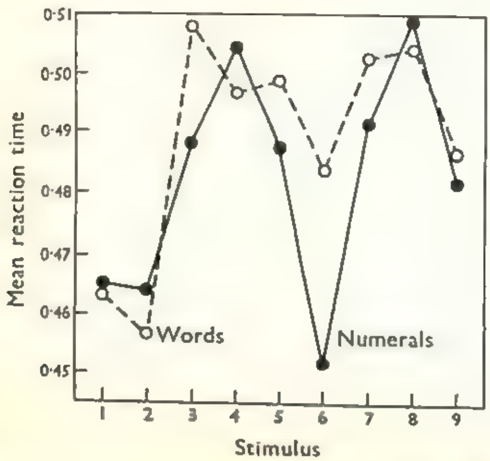
The subjects were 4 men and 14 women students from Introductory Psychology Classes at the University of Wisconsin.

Results

Under the instructions the subjects made few errors; none for digits, one for spelled numbers, and 12 for arithmetic. The data were analysed in terms of each subject's

median RT to each type of stimulus. All statistics were computed from these medians, so that each subject contributed one score for each digit, spelled number, correct or incorrect addition or subtraction, and the data were analysed as randomized blocks, subjects being the blocks, and order of treatment being ignored.

FIGURE 1



Mean reaction time to digits and spelled numbers (words).

In Figure 1 the median RTs for each digit and spelled number are averaged over subjects. The standard errors of these means ranged from 0·012 to 0·016, and the average values tend to be slightly higher than those reported earlier (e.g. by Forrin and Morin, 1966). The data of Figure 1 show that RTs for digits tend to parallel those for spelled numbers, being fastest for 1, 2, 6, and 9, and these differences are statistically significant when tested by an overall analysis of variance, $F\ (8,136) = 2 \cdot 83$, $p < 0 \cdot 01$ for digits and $F\ (8,136) = 2 \cdot 18$, $p < 0 \cdot 05$ for spelled numbers, with subjects $\times$ stimulus type providing error mean squares of 0·0024 and 0·0028 sec. respectively.

We have no convincing explanation for the significant variation in RT to the different numbers. The fastest RTs were to the numbers 1, 2, 6, and 9, and there is a high correlation between RT and length of the spelled number-word, regardless of whether the number was presented as a digit or a spelled word. The differences in RT could arise because of the stimulus or stimulus processing of the subjects, or they could arise because of response properties. If the latter is the case, differential effectiveness of the speech sounds in actuating the voice key would seem to be ruled out, because of (1) the high sensitivity of the voice key and (2) the fact that the sound spectrograms of the different responses seemed not to be correlated with the RTs. For example, the initial portion of the spectrograms for "six" and "seven" are virtually identical, but the RTs differ considerably. We have no data to indicate whether or not there is a difference in the time required to organize and emit the different numbers as responses. Thus, finally, we can only conclude that the numbers do not form a homogeneous class as RT stimuli and responses.

In Table I are given the mean RTs, averaged across subject's medians, for the four classes of arithmetic problems. The analysis of variance treats the experiment as factorial in problem type and correctness in randomized blocks (subjects) and is summarized in the lower half of the table. The error term is obtained by pooling the block interactions. All effects, problem type, correctness, and their interaction, are statistically significant, but subsequent t-tests showed that the only significant differences between means arose because of the lower RT to correct addition problems when compared with the essentially equal RTs to incorrect additions, $t\ (51) = 5 \cdot 91$; and correct, $t\ (51) = 5 \cdot 27$, and incorrect subtractions, $t\ (51) = 6 \cdot 67$. Otherwise the largest t was 1·40. Thus RT or "classification time" for additions was lengthened by incorrectness, but incorrectness did not delay classification of subtractions as correct or incorrect.

Although the average speed of adding and subtracting was of the same order of magnitude as that reported by Blackburn (Fitts, 1964, p. 266), the processing of addition and subtraction evidently differed. Any explanation is somewhat conjectural, but it

TABLE I

MEAN RT IN SECONDS FOR CLASSIFYING CORRECT AND INCORRECT ARITHMETIC PROBLEMS

Problem type	*Mean* Correct	*Mean* Incorrect
Addition	0·896	1·061
Subtraction	1·043	1·082

SUMMARY OF ANALYSIS OF VARIANCE

Source of variation	*d.f.*	*Mean square*	*F*
Problem type	1	0·129	18·43
Correctness	1	0·186	26·57
Interaction	1	0·072	10·29
Between Ss	17	0·095	13·57
Error	51	0·007	

appeared that the subjects, knowing the addition table better than the subtraction table, "recognized" correct additions but had actually to carry out some calculation in order to classify the incorrect additions and the correct and incorrect subtractions.

EXPERIMENT II

Method

The basic procedures and apparatus for recording the verbal classification times to true and false sentences were the same as for the arithmetic problems, except that the stimuli were presented by means of a Gerbrands Model T1B Tachistoscope with a pre-exposure and exposure luminance of 5 ft. candles. The words were printed in India ink with a Leroy lettering set. Each letter was ¼ in. in height and the viewing distance was 2 ft. Each subject had to classify thirty sentences, half of which were true and half of which were false. The sentences were short, such as: "BEES MAKE HONEY" or "HORSES CAN FLY." At the conclusion of the RT procedure, all subjects were required to fill out a rating sheet on 120 sentences, including the 30 sentences to which they had been exposed. The rating sheet classified each sentence on a five-point scale ranging from "definitely true" to "definitely false." The correctness classification of a sentence was considered "uncertain" if two or more of the 18 subjects failed to rate it as "definitely true" or "definitely false."*

Eleven women and seven men students from Introductory Psychology classes were used as subjects in this experiment. The data from two additional subjects were discarded because of equipment failure.

Results and discussion

Under the RT instructions about 10 per cent. of both the true and the false sentences were misclassified. RTs on erroneous classifications were not used. Again each subject

* The experimenters' intent was that all sentences be definitely true or false. In fact, subjects rated only 37 of the 60 "true" statements as definitely true and 31 of the 60 "false" sentences as definitely false. After the fact, the bases of the uncertainty of the subjects were sometimes obvious, sometimes amusing. Examples of dubious "false" statements are: BALBOA DISCOVERED AMERICA, ELEPHANTS ARE TINY, THERE IS A SANTA CLAUS. Examples of dubious "true" statements are: FIRE BURNS, WOOD CAN FLOAT, RAZORS ARE SHARP.

provided four median scores corresponding to the four cells in the upper part of Table II, that were used in the analysis. The means across subjects varied from 1·462 sec. for uncertainly false sentences to 1·258 sec. for uncertainly true sentences as shown in Table II. The statistical analysis was the same as with the arithmetic problems, and as with the arithmetic problems, all sources of variation produced significant *F*s. A series of *t*-tests performed on the different scores between the various stimuli showed that only the uncertainly false stimuli took significantly longer to classify than the certainly false, $t(51) = 3{\cdot}96$; certainly true, $t(51) = 4{\cdot}63$; and uncertainly true sentences, $t(51) = 9{\cdot}11$. Otherwise the largest t was 1·09. The slow performance on the uncertainly false sentences, then, was the basis for all the significant *F*s for sentence type, certainty and their interaction in Table II.

TABLE II

Mean RT in Seconds for Classifying Short Sentences as True or False

Uncertainty of classification	*Mean* "*True*"	*Mean* "*False*"
Uncertain	1·258	1·462
Certain	1·267	1·295

Summary of Analysis of Variance

Source of variation	*d.f.*	*Mean square*	*F*
Certainty	1	0·113	7·06
Correctness	1	0·243	15·19
Interaction	1	0·139	8·69
Between *S*s	17	0·194	12·12
Error	51	0·016	

No interpretation of these data will be offered, but we note that uncertainty degraded performance on false sentences more than on true sentences, and although Gallenbeck and Smith (1950) reported that decision time increases with uncertainty of belief or conviction, it would appear that the subject's information processing problem becomes further complicated when statements are presented in false or negative form. In a somewhat similar context, Wason (1959, p. 107) has referred to the "apparent difficulty of dealing with negative information."

This research was supported in part by a grant from the U.S. Public Health Service to the second author, MH 06792. The authors wish to thank Professor Maurice E. Hershenson for providing the Scientific Prototype tachistoscope used in Experiment I.

References

Fitts, P. M. (1964). Perceptual-motor skill learning. In Melton, A. W. (Ed.), *Categories of Human Learning*. New York: Academic Press.

Fitts, P. M., Peterson, J. F., and Wolpe, G. (1963). Cognitive aspects of information processing: II. Adjustments to stimulus redundancy. *J. exp. Psychol.*, **65**, 423–32.

Fitts, P. M., and Switzer, G. (1962). Cognitive aspects of information processing: I. The familiarity of S-R sets and subsets. *J. exp. Psychol.*, **63**, 321–9.

Forrin, B., and Morin, R. E. (1966). Effect of contextual associations upon selective reaction time in a numeral-naming task. *J. exp. Psychol.*, **71**, 40–6.

Gallenbeck, C., and Smith, K.U. (1950). Systematic formulation and experimental analysis of the phenomena of thinking and belief. *J. exp. Psychol.*, **40**, 74–80.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11–26.
Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.
Morin, R. E., Konick, A., Troxell, N., and McPherson, S. (1965). Information and reaction time for "naming" response. *J. exp. Psychol.*, **70**, 309–14.
Thomas, H. B. G. (1963). Communication theory and the constellation hypothesis of calculation. *Quart. J. exp. Psychol.*, **15**, 173–191.
Wason, P. C. (1959). The processing of positive and negative information. *Quart. J. exp. Psychol.*, **11**, 92–107.

Manuscript received 6th October, 1966.

APPARATUS

ACCURATE DELAYS FOR AUDITORY FEEDBACK EXPERIMENTS

BY

A. W. F. HUGGINS

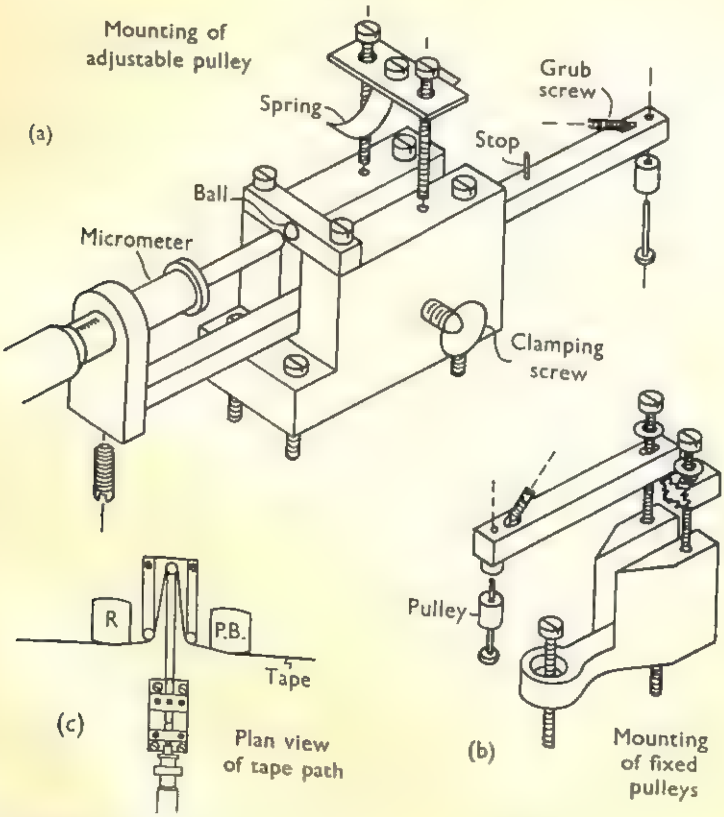
From the Department of Phonetics, University College, London

This note describes a way of modifying a tape recorder for producing accurately controllable delays for experiments with delayed auditory feedback. Any value of delay from 80 millisec. to 1·2 sec. can be obtained to the nearest millisec., and the range could be extended by some minor changes. The delay is continuously monitored on a digital electronic timer.

Producing the delay

The delay is varied by changing the length of the path followed by the tape between the record and playback heads of the recorder, and is thus based on the same principles as the adapter described by Tiffany, Hanley, and Sutherland (1954). This method has the following advantages: an existing tape recorder is used, thus minimizing the amount of equipment to be built, and the record and playback heads do not have to be moved; a complete record is produced, at standard tape speeds, of the speech disturbances induced by the delay; and normal use of the tape recorder is not affected.

FIGURE 1



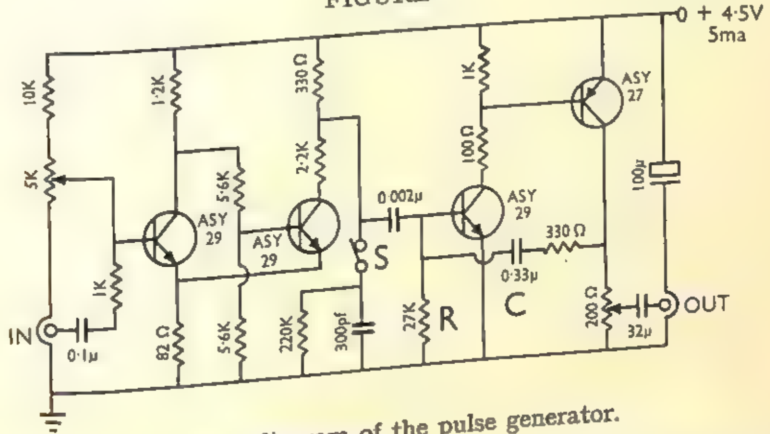
Partly exploded diagram (not to scale) of the component parts of the modification, and a plan view of the modified tape path.

The tape recorder used was a Brenell model STB 1, which is a half-track stereophonic recorder, with one record and two playback heads, and tape speeds of 15, 7½, 3¾ and 1⅞ i.p.s. The "upstream" playback head gap is 1¼ in. from the record head gap, giving a minimum delay of about 80 millisec. at 15 i.p.s. Smaller values of delay could probably be obtained by fitting a larger capstan, thus increasing the tape speed.

The delay is increased by means of three pulleys between the record and playback heads. Two of the pulleys are mounted on rigid arms on a block bolted to the tape recorder base plate between the heads (Fig. 1*b*). The third pulley is mounted on a sliding arm, and can be brought up from the back of the tape to deflect the tape path between the fixed pulleys (Fig. 1*c*). The sliding arm is held by a spring and a clamping screw in a second block, that is bolted to the base plate on the far side of the pressure-pad mechanism from the tape. When the sliding arm is drawn back from the tape, its travel is checked by a stop that prevents damage to the pulley. In this position, none of the pulleys is in contact with the tape, and the tape can be threaded and rewound, and the recorder can be used exactly as if unmodified. The sliding arm can be slid forward until a micrometer, mounted on the arm, touches a ball, fixed to the block holding the slider (see Fig. 1*a*). Thus the setting of the micrometer controls the position of the movable pulley, and thus the delay, and yet the whole slider can be drawn back, out of contact with the tape, without altering the setting of the micrometer.

Since a 1-in. micrometer was used, the tape path can be lengthened by about 2 in. to 3¼ in., giving a delay of about 220 millisec. at 15 i.p.s. If the second, "downstream" playback head is used, the tape path is lengthened by a further 1¼ in., giving a maximum delay, at 15 i.p.s., of 300 millisec. Longer delays are obtained by using the slower playback speeds, the 3¾ i.p.s. speed giving a maximum delay of 1·2 sec. A larger range of delays could be obtained by using a 2-in. micrometer, but problems might then be encountered with the vertical alignment of the tape. For some purposes, it might also be possible to use the 1⅞ i.p.s. speed, though the quality of the reproduction deteriorates, and flutter might become objectionable.

FIGURE 2



Circuit diagram of the pulse generator.

The hardest problem is making the pulleys, since they must be very small to fit three-in-line between record and playback heads. The pulleys were ⅛ in. in diameter, and were made from PTFE. The pins they run on were $\frac{1}{16}$ in. silver steel. The difficulties arise because departures from roundness of more than about 0·0002 in. cannot be tolerated, and each pulley must be a friction-free and wobble-free fit on its pin. However, acceptable pulleys were made on an ordinary metal-lathe, as follows. The PTFE is held in a collet while the central bore is centred and drilled with a No. 53 drill. Then the O.D. is ground (or turned, taking small cuts with a sharp rounded tool) with about 3 in. of PTFE projecting from the collet, and the "floating" end supported by a steel pin in the bore, the steel pin being held in a pinchuck in the tailstock. Finally the I.D. is reamed to $\frac{1}{16}$ in. It is also important that the ends of the pulley be flat, and normal to the bore. With pulleys made in this way, and using (Scotch) instrumentation tape, it was possible to reduce the wow and flutter, by adjusting the feed spool holdback tension, until it was inaudible, even in a 1 kcs. pure tone. The requirements for speech are not so stringent.

Monitoring the delay

The Brenell STB 1 is a stereophonic recorder, and in its present application, only one track is used for delaying the speech (or other) signal. The second track is used for monitoring the delay. A pulse is recorded on the second track, and when it is played back, after the delay, it is used to trigger the pulse generator, so that a second pulse is recorded, and so on. The result is a train of pulses on the second track of the tape, whose repetition rate is the reciprocal of the delay. If a second output from the pulse generator is fed to an electronic timer (e.g. Venner TSA 3314, Advance TC2A, Hewlett-Packard 3734A), a digital readout of the delay in millisec. can be obtained. The micrometer adjustment, together with the continuous digital monitoring, makes possible very accurate adjustment of the delay.

The circuit of a suitable pulse generator is shown in Figure 2. It consists of a Schmitt trigger for detecting and reshaping the pulses played back from the recorder, followed by a monostable multivibrator that produces the pulses for recording. The pulse train is started when switch S is closed with the recorder in the record mode. The circuit shown delivered pulses lasting about one millisec., but the pulse duration depends linearly on the value of condenser C. The monostable was designed to have a recovery time of about 30 millisec., to prevent double pulses being recorded, but the recovery time can be reduced by reducing the value of resistor R, if desired. (For a useful introduction to transistor pulse circuits, see Towers, 1965.)

References

Tiffany, W. R., Hanley, C. N., and Sutherland, L. C. (1954). A simple mechanical adapter for variable sidetone delay. *J. Speech Dis.*, **19**, 504–6.

Towers, T. D. (1965). *Elements of Transistor Pulse Circuits*. London: Iliffe, and Princeton, New Jersey: Van Nostrand.

Manuscript received 8th August, 1966.

BOOK REVIEWS

Contemporary Psychology Experiments: Adaptations for Laboratory. By John Jung and Joan Bailey. London and New York: Wiley. 1966. Pp. xiii + 155. 23s.

The title describes the contents well. Among others, two important advantages are claimed for this book. First, that the topics of the 15 experiments being drawn from recent publications and dealing with current issues, they will be of immediate interest to students undertaking them. Secondly the work is designed to be exploratory rather than confirmatory, and this is expected to give students a greater feeling of involvement, since they will to some extent be breaking fresh ground and their own results will contribute to the clarification of the underlying issues. Both features are valuable and the book deserves careful attention from those who run laboratory courses.

On the basis of experience I see other points as calling for comment.

(1) The 15 "experiments" in this book are all connected with current issues, but otherwise they are independent and they could be done in any order. Some operators may see this as a convenience, but there is great value in a coherent set of "experiments" where what has been learnt is built on and where the student can feel that he is making progress in outlook and ideas as well as in technique.

(2) The form in which data are to be reported and the questions to be considered are very precisely specified. Uniformity has been sought in order to ensure comparability between results obtained by a number of "experimenters." My experience has been that the more precise the specification the less likely students are to keep their eyes open and observe what is happening.

(3) Only in some of the "experiments" is the "experimenter" called on to ask questions of the Subject. In this way a fruitful source of information is neglected. It is always of value to find out what the Subject has to say.

(4) Statistical treatment is by way of hypothesis and verification, a route which tends to inculcate rule of thumb methods and offers little encouragement to observers to think what it is they are collecting information for and how to get it.

If this book is used for beginners it is of enormous importance that they should sharpen their statistical eyesight by the use of descriptive methods. T. L. Kelley used to say of correlation that he always went back to the scatter diagram because all the information is there, while the correlation coefficient only summarises the information in the table relevant to the concept of linear correlation. Furthermore the basic problem in many psychological situations is not to obtain numerical data to play with, but to maintain a clear understanding of how the numerical data are related to the psychological concepts involved.

(5) "Variations in procedure from one E to another could invalidate the entire experiment, since there could be no clear basis for interpreting the combined class data. Hence it is extremely important that you follow in detail all instructions given in the workbook and by your instructor."

In its proper place this advice is sound, but it implies a blindness to the realities of the laboratory. Instructions sufficiently clear to enable a group of students to start working nearly always reveal differences in emphasis or interpretation from one student to another. This is a most fruitful source of fresh information about situations and about people and should never be despised or neglected. If students are to carry out procedures precisely they must be trained in those procedures.

The combined effect of these five comments is that the book reviewed is a useful but a very rigid framework for a laboratory course. Its rigidity will train students to carry out precise operations to the best of their ability. Prospective users must consider the complementary effect, that students will receive no encouragement to be observant, or to try to understand the people they deal with, the nature of human responses and feelings or, and this is perhaps the most serious failing, to learn from their mistakes.

It would of course be possible for anyone who so wished to take the situations and to use them with their own modifications, but to do this would be to build on the book as a foundation, not to use it as it is.

B. BABINGTON SMITH

A Textbook of Psychology. Second Edition. By D. O. Hebb. Philadelphia and London: W. B. Saunders & Co. 1966. Pp. xvi + 353. £2.

The first edition of this well-known text was reviewed in this *Journal*, 1959, **11**, 189. This edition has been extensively re-organized but retains the general approach and lively

style of its predecessor. There is now less emphasis on early learning and a more cautious appraisal of the relevant evidence, both clinical and experimental. On the other hand, there is much more about adult learning and intelligence, too, gets a fuller treatment, including a section on Piaget. The sections on the nervous system have been slightly expanded and brought up to date. A glossary has been added—useful in many ways but with some unfortunate entries. "Disgust," for instance, is defined as "arousal accompanied by mediating processes tending to produce avoidance of the exciting object." Surely this is just the kind of thing that Dr. Hebb has hitherto been so successful in *not* writing! Glossaries should convert jargon into English, not *vice versa*.

O. L. ZANGWILL.

Behaviour Therapy Techniques: A Guide to the Treatment of Neuroses. By Joseph Wolpe and Arnold A. Lazarus. Oxford and London: Pergamon Press. 1966. Pp. ix + 198. 21s.

This book consists mainly of a detailed description of behaviour therapy techniques. Particular attention is paid to the methods of "reciprocal inhibition" popularized by Wolpe in which an action or thought inimicable to the neurotic symptom is practised by the patient to help overcome his difficulties. Other techniques involve "teaching," explanation, suggestion, persuasion, drug abreaction, conditioned avoidance techniques and environmental manipulation.

No-one would deny that behaviour therapy is a useful addition to psychiatric treatment methods for selected patients. Many, including the reviewer, would suggest that the presentation by Wolpe and Lazarus is oversimplified and that some of the claims made (e.g. that "behaviour therapy is effective in all neuroses") are simply not true. The approach to basic theoretical problems of conditioning and learning is naive; the diagnostic interviewing (examples are quoted in detail) is crude; the techniques themselves, as suggested above, are blunderbuss in their variety; some of the moral issues are frankly dubious (e.g. active encouragement, if necessary, of extra-marital sexual relations). These and other features suggest therapeutic enthusiasm without either wisdom or balance. The chapter on "results" is selective and brushes aside or ignores investigations which are less than wholly favourable. From an evaluative point of view the chapter is worthless. The frequent references to psychoanalytically based schools of psychotherapy show a lack of understanding of both psychodynamic theory and method.

As an elementary handbook of behaviouristic techniques in psychiatric practice, this is a useful book. As a contribution to a comprehensive and coherent theory of behaviour it is valueless.

SIDNEY CROWN

Motivated Learning: a Developmental Study from Birth to the Senium. By Miriam E. Hebron. London: Methuen. Methuen's Manuals of Modern Psychology. 1966. Pp. 264. 35s.

This book represents an attempt to trace the development of human motives through the life span using two basic assumptions: first that a subject will seek to maintain an optimal state of arousal, avoiding extremes of high and low; and second that, although doing this will determine what is attended to and therefore what is learnt, the manner of maintaining optimal arousal will itself depend on previous learning, so that motivation and learning will each develop the other. The first chapter outlines evidence on the physiological mechanisms involved in arousal and learning, and the remaining chapters consider the development of skills and motives at various periods of life from the first two years up to old age.

The aim is important and timely, and the wealth of ideas and penetrating insights cannot fail to stimulate the reader even if, as is likely, he is inclined to question many of the views advanced. Perusal of the book is unfortunately difficult owing to the manner of presentation. Evidence is discussed minutely without always being clearly stated, while the steps leading to the author's own conclusions are often not set out in detail, with the result that the argument frequently appears at once discursive and dogmatic. Many of the physiological suggestions seem reckless in that, although they are brave, they do not really contribute anything essential to the thesis being developed. There is also a great deal of difficult jargon and a number of neologisms and errors of grammar and spelling.

This reviewer does not consider himself competent to judge the scholarship of Chapters 2–7 on childhood and adolescence, but it appears to be much more thorough than that of

chapters 8 and 9 which deal respectively with maturity and old age. Little of the substantial volume of work published during the last 12 years is mentioned in Chapter 8, and in Chapter 9 it is completely ignored.

These faults should, perhaps, be judged leniently when the immense scope of the work is remembered. To weld physiological and experimental findings on arousal and attention together with Freudian concepts and developmental evidence extending over the whole life span is indeed a formidable task, and Dr. Hebron deserves our gratitude for having attempted it at all. We may hope, however, that having come thus far she may one day take the further step, which someone will have to take sometime, of cutting through the tangled concepts and fantastic terminology of previous writers to a straightforward, coherent and lucid statement. In short, this reviewer would say to Dr. Hebron what a reader of his own first book said to him: "I look forward to reading your *next* book."

A. T. WELFORD.

Human Brain and Psychological Processes. By A. R. Luria. Translated by Basil Haigh. New York and London: Harper & Row. 1966. Pp. xix + 587. £5 10s., $14.50.

Higher Cortical Functions in Man. By A. R. Luria. Translated by Basil Haigh. London: Tavistock Publications. 1966. Pp. xvi + 513. £6 6s.

The first volume brings together ten papers on neuropsychological topics written by Professor Luria over the past 30 years, several of which have been specially revised and brought up to date for publication here. Among the problems dealt with are cerebral localization of function, the cortical organization of movement and the role of language in the regulation of behaviour. Although the main emphasis is upon the effects of brain lesions, Professor Luria has a wide acquaintance with general psychological issues, more especially in the developmental field, and is concerned throughout to use the data of pathology as an aid to our understanding of normal behaviour. This book could be read with profit by any psychology student sufficiently intrepid to venture beyond the fashionable texts and topics of today.

The second volume really is a work of major proportions, first published in Russia in 1962. It embodies a full and carefully presented account of the main types of psychological disturbance resulting from local lesions of the brain, with a wealth of illustrative clinical material. Appropriate methods of clinical examination, many of which we owe to the author himself, are described at considerable length. Although the lack of quantitative data may disturb some psychologists, anyone who once read and learned from the writings of such men as Head, Goldstein or von Monakow will find this volume in the highest intellectual traditions of European neurology.

Prefaces to both volumes have been contributed by Dr. Lukas Teuber and Dr. Karl Pribram. As usual, Dr. Basil Haigh's translations are altogether excellent.

O. L. ZANGWILL.

Productive Thinking. By Max Wertheimer. Edited by Michael Wertheimer. Social Science Paperbacks. London: Tavistock Publications. 1966. Pp. xvi + 302. 30s.

A reissue of Max Wertheimer's famous book first published in 1945, shortly after the author's death. This edition contains three additional chapters and several new appendices based on material found among the author's notes and papers. Though providing interesting further illustrations of Wertheimer's thinking, this new material does not add anything radically new to the principles embodied in the book. None the less, admirers of Wertheimer will be indebted to the Editor for his devotion in putting together this new and enlarged edition.

O. L. ZANGWILL.

Problem Solving: Research, Method and Theory. Edited by B. Kleinmuntz. London and New York: Wiley. 1966. Pp. ix + 406. 53s.

This volume reports the first of a series of annual symposia on cognition to be held for the next several years at the Carnegie Institute of Technology. Thirteen distinguished participants provide an introduction, seven papers, three contributions from discussants, and an epilogue. An attempt is made to give coherence to the whole by contrasting the approaches of "information processors" and "behaviourists." However, the contributions are not intimately related to each other and, as with so many symposia, it is the occasional individual paper or comment which attracts attention. This makes it difficult to say briefly which segments are likely to interest which readers. So far as this reviewer

is concerned, the interest lies in these segments of the book which report new work, notably, some brilliant experiments by J. R. Haynes on the organization of memory in problem solving, a sophisticated study of "algebra word" problems by J. M. Paige and H. A. Simon, and some fresh data from A. D. de Groot on competence in chess.

I. M. L. HUNTER.

Human Conceptual Behavior. By Lyle E. Bourne, Jr. Boston: Allyn and Bacon. 1966. Pp. viii + 139. $2.50.

This workmanlike little volume is part of a new paperbound series entitled "Contemporary Topics in Experimental Psychology." The editors are James Deese and Leo Postman, and each volume aims to introduce undergraduates to ways in which some set of psychological issues can be explored by appeal to experimental data. The issues in this volume arise from those kinds of laboratory task which have come to be known as experiments in concept formation. The book has three roughly equal parts. The first part describes what is meant by "concept," some different forms of concept experiment, and the main ways in which these experiments have been interpreted. The second part pursues the experimental analysis of several task variables, and the third part continues this analysis in relation to subject variables. Within his self-imposed limits, the author skillfully raises clear questions and attempts to answer them by considering relevant experiments of recent date. There is a good bibliography. If other volumes in the series are as competent as this, the series will be useful for undergraduates and also for those of us who find it increasingly difficult to keep abreast of all that is happening in experimental psychology.

I. M. L. HUNTER.

Attraction and Hostility. An Experimental Analysis of Interpersonal and Self Evaluation. By Albert Pepitone. London: Tavistock Publications. 1966. Pp. ix + 234. 30s.

It would be difficult to disagree with the first point made in this book: many theories in social psychology are too general and amorphous to prove their usefulness at the level of specific and unambiguous experimental predictions. This is particularly true, Pepitone goes on to argue, in the area of research with which he is concerned: hostility and attraction in interpersonal relations. A rapid (if sometimes patchy) demonstration follows. The author considers in turn the major need-frustration, need-satisfaction and cognitive balance theories, and finds them all wanting. Each of these theories engendered experiments which leave room for alternative and equally plausible explanations of the results.

The choice of the title of the book is unfortunate. Its width of scope does Pepitone the same sort of disservice that causes him to commiserate with the fate of other theorists. The book is not concerned with a cosmic view of attraction and hostility. More modestly, its aim is to consider a few specific aspects of these phenomena, and within these limits it presents a useful experimental analysis. Many of the measures used in experiments on attraction and hostility are essentially no more than sets of judgements which consist of the subject's evaluations of himself and others on certain specified dimensions. Pepitone's starting point is the hypothesis that, because of a need for "cognitive validation" of the individual's assumptions which underlie his interpersonal relations, these evaluations tend to be constantly revised in the ever-changing social situations. These revisions will take place whenever a discrepancy exists between events that *should* have occurred had the initial evaluations been correct, and events that *do* occur. The experimental paradigm follows directly: it consists of the creation of such discrepancies while attempting to hold constant those determinants of changes in evaluation which are presumably not derived from the need for cognitive validation—e.g. threats to status or security, or the structure of rewards in a competitive situation.

Many of the predictions presented in the book and experimentally confirmed could have been just as well made using sound horse sense; though one need not necessarily share the view that close agreement with common sense is a catastrophe to be avoided at all cost in psychological research. Pepitone manages a simple formalization of a highly confusing aspect of social behaviour. And some of the inferences he draws from his analysis are by no means self-evident; e.g. that more projection will occur in people who have low self-esteem and receive favourable information about themselves than in people with high self-esteem who are placed in the same position.

In his concluding chapter Pepitone provides a piece of sane self-evaluation: "The step has been a small one, and much tracking still needs to be done." Much of this tracking

will have to take place in tangles of variables which, when they pertain to human social behaviour, have a way of stubbornly resisting many of the controlling techniques which are possible in simpler situations. HENRI TAJFEL.

Ability Structure and Subgroups in Mental Retardation. By John Clausen. London: Macmillan. Washington: Spartan Books. 1966. Pp. viii + 208. 80s.

This is a research report. The work is extensive and is carried out carefully, and the presentation is clear and thoughtful. Different age groups of subnormal children were tested on a large variety of tasks and compared with 8 to 10-year old normals. The conclusions were that the subnormal subjects performed generally better on perceptual than on motor tasks, and that there was a relatively greater impairment of simple than of complex motor functions. There was no marked difference between normals and subnormals on measures of eye dominance, two point differentiation, brightness differentiation, mirror drawing, the extent of the Müller-Lyer illusion and anticipation. Those tests which discriminated best between normal and subnormal children included pure tone threshold, simple RT and fine as well as gross motor co-ordination.

Within the subnormal group, muscular strength, static equilibrium and decrease of the size-weight illusion developed from age eight up to adulthood, while performance of motor speed, motor manipulation, dynamic equilibrium and complex perceptual tasks developed up to the age of 15 and thereafter declined. Kinaesthetic and cutaneous space discrimination declined from age 8 onwards. Laterality and performance on some perceptual tasks requiring judgement of quantitative relations showed no development with age, but also no difference between normals and subnormals. On the other hand, no developmental trends but lower levels than normal were obtained on perceptual tasks requiring qualitative judgements and on pure tone threshold.

Analysis of performance patterns according to neurological, EEG or etiological variables did not prove very productive. Such efforts might have been more successful if more determined attempts had been made to link the findings with those obtained in neuropsychology. Thus, the very high pure tone threshold in the subnormals contrasted with a near normal speech threshold. The author explains this in terms of lack of motivation, attention and arousal. But suggestions arising from neuropsychological work such as Milner's, that right-sided brain lesions may interfere with non-verbal auditory perception while lesions of the dominant hemisphere are more usually associated with verbal deficits, might have been taken into account.

Maybe it would be advisable to give explanatory theories in terms of reticular system dysfunction a temporary rest. They have been rather overworked, and have tended to become so general as to appear almost meaningless. Moreover, in the context of the results reported in this book, they are far from convincing. Thus, while the results of this investigation are very interesting and the findings stimulating, a more sophisticated theoretical approach might have improved their interpretation. BEATE HERMELIN.

The Biologic Basis of Schizophrenia. By Jon L. Karlsson. Springfield, Illinois: Thomas. 1966. Pp. xi + 77. $4.75.

This monograph presents a brief account of the aetiology of schizophrenia with particular attention to the genetic aspects. The psychological and biochemical aspects receive mention (with gaps, e.g. Wynne and Singer's work). The genetical account reviews modern concepts of "recessive" and "dominant" (that helped clarify the reviewer's hitherto confused ideas on this subject). An extensive schizophrenic family tree from Iceland is presented in great detail—but without carrying much conviction. A study of the incidence of schizophrenia in the children of a schizophrenic parent, reared in foster homes gives better support to the genetic theory. Lastly the author presents an ingenious and speculative theory trying to link a possible genetic fault with a possible biochemical disorder underlying the disease. The price, however, is excessive for such scant returns. J. R. SMYTHIES.

Psychology for Psychiatrists. By C. G. Costello. London: Commonwealth and University Library, Pergamon Press. 1966. Pp. 328. 25s.

Any young psychiatrists buying this book as a crib for the DPM, or older ones thinking that it is going to put in a nutshell what psychologists know, are, I am afraid, going to be disappointed. This is not psychology as most psychiatrists probably think of it today, but psychology as modern psychologists would like them to think of it.

The book is essentially a résumé of research which has been carried out in a number of fields having some bearing on psychiatric problems It inevitably raises more questions than it answers, but by dealing with them in operational terms, indicates that their answer is within reach. The author wisely keeps clear of the stumbling blocks of diagnostic categories and confines himself instead to such topics as Arousal, Anxiety, Stress and Roles, quoting nearly 900 investigations into these and allied subjects.

Every psychologist should encourage his psychiatric colleagues to read the book. Although written in a straightforward style and refreshingly free from jargon, the mere weight of material makes it heavy going, but we should be extremely grateful to Dr. Costello for collecting it all together for use, and for presenting it in such an unbiased way.

MOYRA WILLIAMS.

Science and Theory of Psychoanalysis. Edited by Irwin G. Sarason. London: Van Nostrand. Insight Book No. 26. 1965. Pp. xiii + 205. 14s.

Professor Sarason has brought together a collection of papers that makes an interesting and useful volume. The papers fall into three parts. The first begins with an (abbreviated) version of Freud's 1913 paper on "The Claims of Psycho-analysis to Scientific Interest"; and it contains papers by Frieda Fromm-Reichmann, Ramzy and Szasz. Part two contains two papers—by Robert W. White and Urie Bronfenbrenner on certain developments of psychoanalytic formulations. Part three contains five papers offering different evaluations of psycho-analysis. It begins with criticisms from Knight Dunlap and Skinner; it continues with a different type of assessment by Joseph Nuttin and a defence by J. F. Brown; and it ends with a review of the scientific status of analysis by Ernest Hilgard.

It is useful to have these interesting papers put together in a handy volume. The harassed teacher and puzzled student will find this volume of value in enabling them to put their hands quickly on a variety of different contributions and points of view on psycho-analysis. However, it could also be argued that the volume has an important limitation. It does not include any paper that explicitly attempts to uncover the sources of the dispute about the status of psycho-analytic theory. The volume does not attempt to explain how it is that distinguished psychologists and others can come to such very different conclusions about the nature of psycho-analysis as they do in this volume itself. To Dunlap, for example, "there was, however, nothing in Freud's system which had not appeared in superstitions which were common several centuries before the beginning of the Christian era." To J. F. Brown, on the other hand, the "theoretical conceptions of psycho-analysis are the most adequate we have in accounting for the sources and distribution of energy in the integrated behaviour of the organism as a whole." In the face of such unresolved conflicts of opinion the ordinary student may be tempted to say, on closing the volume, that he is still almost as confused about the nature of psycho-analysis as he was when he opened it.

But this response to the volume would not be wholly just. For some papers—for example, Nuttin's and Hilgard's—do throw much needed light on some of the sources of the trouble. Thus, Hilgard points out, in effect, how psycho-analytic theory can be looked upon as providing models of different aspects of human functioning; and thereby he shows that, if we do *not* look upon it in this way, we may be tempted into Skinner-like criticisms of psycho-analysis. Again, when he speaks of the validation of psycho-analytic propositions, he speaks of it as involving a "reconstruction" of analysis, from which a "better science will emerge." In so doing, he is also pointing out that, if we do *not* see how the validation of psycho-analytic propositions involves their reconstruction, we may be tempted to maintain, simpliciter, that they cannot be tested at all, and hence that belief in them is just superstition. So there is much more in this volume than the casual reader may suppose; and some of the papers deserve careful attention.

B. A. FARRELL.

Aspects of Learning and Memory. Edited by D. Richter. London: Heinemann. 1966. Pp. x + 182. 35s.

Enough books purporting, as does this one, to describe recent advances in memory research are now making an appearance to give the conscientious reader an all too direct experience of the reality of retroactive and proactive inhibition. In the present case the material is internally so variable that not much interference is likely to occur between its chapters. Indeed a few of the chapters do not easily interfere with the retention of

other works even in the same field of specialization. According to the book jacket "It is written by a group of scientists working in different fields of research and each chapter gives a different viewpoint. The first three, written by a psychologist [M. Metcalfe], a clinical neurologist [W. Ritchie Russell with Freda Newcombe] and a pathologist [J. B. Brierley], deal with various aspects of memory in man; the remaining four chapters relate to basic aspects and pursue the problem further into the fields of physiology, biochemistry and animal behaviour." These latter four authors are Sir John Gaddum, D. Richter, B. S. Meldrum and I. Steele Russell.

Not only is the work variable in subject matter, but also in the thoroughness of treatment and in application of critical judgement. The chapters by W. R. Russell and Freda Newcombe, and by Brierley, are useful and concise summaries of some of the salient features and causes of memory disorders seen clinically. Even these, particularly the former chapter, contain some puzzling assertions, such as the argument that disorders caused by injury to the limbic system could not, in effect be a failure of retrieval "as far as learning new material is concerned" (p. 17). On the whole the four papers which relate to physiology, biochemistry, and animal behaviour are much more detailed and closely argued. Richter's chapter on biochemical aspects of memory is level-headed and cautious, and contains a helpful critique of some features of theories of molecular coding of memory. Richter seems to favour the more traditional view of an alteration in synaptic transmission to account for learning, a view also endorsed by Meldrum in his chapter on electrical signals and cellular mechanisms of learning. Strong evidence favouring such a view is difficult to pin down, but at least a number of speculations are presented here. Meldrum's paper seems unnecesarily jargonistic, but it provides a varied diet and some unusual suggestions. He appears to be somewhat sceptical of others' views and is dogmatic in support of his own suggestions. And both he and Richter seems all too ready to accept Lashley's views on mass action without any attempt to incorporate material more recently available—including material presented elsewhere in the same volume.

We are told in the preface that the chapter by the late Sir John Gaddum is the last of his scientific publications, and there is a suggestion that he was under considerable stress when it was written. The chapter is lively and displays an enormous range of interests and knowledge. But it is, unfortunately, terribly uncritical and will disappoint those whose expectations are based upon Sir John's earlier writings. But at least he himself warned us in his paper that "various assumptions have been made without evidence and the language of psychology has been confused with that of physiology, and fact with fancy." One is not inclined to quarrel with that, but few will not be stimulated by him.

I. Steele Russell's chapter is, to my mind, the most solid and scholarly contribution in this volume, and the most comprehensible to the average experimental psychologist. It contains an excellent review of relevant animal experimental material, including stimulation, lesion, electrophysiological recording, and consolidation types of studies. It can be recommended as a rich source of references. But more noteworthy is the clarity of the distinctions among various types of processes and his insistence that learning is a far from unitary affair. In a chapter of this scope any reviewer is bound to disagree with points of emphasis, and to expose his own biases. I feel that his treatment of the effects of temporal and frontal lobe lesions is too limited and underplayed, while certain other work, such as Gastaut's electrophysiological research is overplayed. Overplayed, also, is the reification of habituation into "redundancy learning." (Not that the idea is without merit, but I feel the distinction between perceptual adaptation and habituation is often quite difficult to make, and few would wish to say that a "negative after-image" is an example of "redundancy learning.") Also, there are one or two minor but misleading errors, such as his statement that Weissman, Chevalier, and King found "permanent memory losses" after E.C.S. (p. 155). In fact, only Chevalier examined the permanence of the losses. But this chapter is solid and valuable, and some of its words of wisdom could be read with benefit by certain of the other contributors.

L. WEISKRANTZ.

The Spiral After-Effect. By Harry C. Holland. London: Pergamon. 1965. Pp. xii + 99. 35s.

The movement after effect (MAE) is one of the most striking and intriguing illusions in psychology. First reported by Aristotle, it was one of the very few effects to be completely misinterpreted by Helmholtz (he incorrectly attributed it to eye movements), and 100 years and over 200 papers later, its origin is still an unsolved mystery. Holland's

monograph is a model of lucidity and thoroughness. It contains a well balanced critical review of the literature from 1825 to 1963–4, with chapters on the influence of Stimulus variables, Programme variables (presentation rates, etc.) and Observer variables: and theoretical explanations of the effect. He makes suggestions for further research, which he facilitates by providing, in an appendix, reproductions of 31 useful spirals. As a good experimentalist, Holland suggests that many of the discrepancies between the findings of different workers almost certainly arise from a failure to standardize experimental conditions. He is quite prepared to criticize individual experiments (e.g. "this present worker has three times failed to replicate this observation with older observers" (p. 9).

The MAE was widely studied during the nineteenth century, and much of the best work on it was done then. During the revival of interest in the effect over the last 20 years, much of this early work has been repeated, often unwittingly and often less skilfully. There has been much recent experimentation both by "pure" experimentalists and by clinicians looking for diagnostic tests of personality and of brain damage. Holland explores the tangled web of experiments designed to evaluate the MAE as a clinical tool, concluding that psychologists "would be unwise to imagine that they are conducting an investigation of brain damage, just as any neurological surgeon would be equally unwise to accept the results of spiral 'tests' as factors in their diagnostic deliberations" (p. 73).

It is curious and sad that an effect which has been studied so much should be understood so little. A mere 4½ pages (out of 99) are enough to describe all the past and present "explanations" of the effect—none of which commands universal agreement. This is only half a page more than Holland's useful suggestions for further research. Oddly, Holland seems content to be agnostic about causes: "The person who reads this monograph in the belief that it will 'explain' the spiral after-effect will be disappointed, for it is based upon a working premise that the main interest of the psychologist is *how* and *under what conditions* the illusion occurs or can be modified. It is not concerned with the 'ultimate cause'—whatever this may mean—but only with its description" (Introduction, p. x). But in spite of Holland's lack of interest in underlying mechanisms, his monograph will now be essential reading for everyone who hopes to understand the MAE.

S. Anstis.

Dynamics of Response. By J. M. Notterman and D. E. Mintz. London and New York: Wiley. 1965. Pp. 277. 70s.

The authors begin with the proposition that any operant must involve some exertion of force over a period of time, and that failure to investigate this aspect of responding has caused a range of experimental and theoretical questions concerning operant behaviour to be neglected. The experiments reported, most of which are published here for the first time, serve to pose and give partial answers to some of these questions. All the experiments use rats trained in a specially designed Skinner box in which force of lever pressing is accurately measured. The novel dependent variables utilized, peak force of response and time integral of force of the response, are shown as being significant in that they can be differentiated independently of rate of response.

Given this the problems attacked centre round the attempt to specify the relations between effort demanded and effort expended. Evidence is presented which suggests that the rat tends not to maximize reinforcement as such, even when effort requirements are low, but to approach a position in which "amount of reinforcement per unit of force" is maximized. This conclusion arises out of a range of studies on the effort requirements for reinforcement in conventional and proportional reinforcement situations. The authors see as central to the understanding of these relations the discriminative processes involved in them. The thesis propounded is that response differentiation represents a special case of stimulus discrimination, one in which the cues involved are cutaneous and proprioceptive. It is backed by a range of experiments including several interesting developments involving reinforcement of one or two "bands" of intensities of response, with and without external cueing and internal feedback, and is extended to original studies on force measurements in fixed ratio and interval schedules as well as to studies on drive and amount and quality of reinforcement.

The authors appear to have aimed mainly at mapping out the field, with the inevitable consequence that their data in some cases are extremely thin. The mapping process is certainly a success, and despite its deficiencies there is sufficient evidence to indicate important consistencies in phenomena. For example, increase in force is observed under unreinforced response conditions, this observation holding under conventional extinction,

within cycles in fixed ratio and fixed interval, and in the process of discrimination of stimuli which signal the availability or unavailability of reinforcement. Secondly the use of force measures as dependent variables and force requirements as independent variables, allows the development of new procedures which have potential value in the examination of problems outside the range of this monograph, and it seems clear that its publication represents a further step in the extension of the range, and consequently the analytic and explanatory power of operant techniques.

C. C. KIERNAN.

Contrary Imaginations: A Psychological Study of the English Schoolboy. By Liam Hudson. London: Methuen. 1966. Pp. vii + 181. 25s.

This book has three outstanding merits. It concerns real psychological issues about relatively mature intellectual functioning; it treats these issues with a constructive honesty which grants their complexity; and it is written in forcefully simple language which is a pleasure to read. These merits need to be stressed because there is much which invites criticism—a needlessly iconoclastic flavour, omissions of several obviously relevant psychological studies, metaphors which are at once vivid and acknowledged to be inadequate, considerable speculative indulgence, an occasional lack of sharp clarity in the questions pursued. These defects will irritate readers in different ways and to different extents. But they are worth bearing for the sake of the book as a whole.

The factual core concerns the performances of several hundred English schoolboys on assorted tasks. The boys are in the senior years of secondary school, and are all intellectually accomplished in that they are likely university entrants. The tasks comprise an intelligence test (AH5), open-ended tests (such as naming as many uses as possible for some specified everyday object), composing a short autobiography and description of interests, freely commenting on controversial statements, expressing opinions about personal qualities, illustrating a verbal phrase with a drawing. So the factual core is substantial. However, the author does not aim to overpower the reader with detailed results. Rather, his aim is more interesting and ambitious. He strives to make overall sense of his data and his impressions by extracting general trends while also insisting on the individuality of each boy who contributes to these trends. To accomplish these laudable aims, he frequently resorts to the powerful but hazardous device of the significant contrast. In particular, he organizes much discussion around the contrast between convergent and divergent casts of mind; and in so doing, he greatly clarifies this dichotomy and divests it of some of the value-laden nonsense which has surrounded it in recent years.

The convergent-divergent contrast emerges from biased accomplishment on two distinguishable kinds of task. One kind is provided by conventional IQ tests which typically require the person to *converge* his efforts towards one correct answer. The other kind of task is provided by open-ended tests which require the person to *diverge* fluently without examining any one line of reasoning in detail. When performances are scored, there are low and positive correlations between the IQ battery and the open-ended battery (and also, note, between sub-tests within the same battery). These imperfect correlations enable Hudson to draw out, from the total sample of boys, two extreme groups; one distinctly biased towards convergent and the other towards divergent accomplishment. Having got these extreme groups, he proceeds to detailed elaborations from which this reviewer would extract three important points. (*a*) This test-score bias is an indicator of a more general predispositional bias which shows up in several contexts, such as the kind of autobiography produced, value judgments, interest in arts subjects as opposed to sciences. (*b*) The contrast is a matter of bias rather than level of accomplishment and, further, most boys do not show strong bias. (*c*) Both types of bias have their weaknesses and strengths and also (a point which needs stressing) a divergent bias is certainly not the same thing as being "creative." These three points are clearly made but, in the book, they are mixed up with various speculative digressions, some highly insightful, some distinctly unsatisfactory. In particular, the author is not at his best when discussing what might be called motivational issues, for example, his sorties concerning rival systems of defence and the genesis of original thinking.

On balance, this is a worthwhile book. To the professional psychologist, it gives much factual work, some careful conceptual analysis, and a refreshingly commonsensical attack on the organization and development of higher-level human accomplishments. To non-psychologists, for whom this book may well be very influential, it gives a readable demonstration that "human faculty" is still an important and challenging topic for empirical enquiry. The author deserves thanks for what he has done and encouragement for his future explorations.

I. M. L. HUNTER.

The Structure of Association in Language and Thought. By James A. Deese. London: Oxford University Press; Johns Hopkin Press. 1966. Pp. xiii + 216. 52s.

Professor Deese in this book heralds a new approach to the theory of word association. He argues that the philosophers have misled us into considering association as a development of stimulus-response pairs. Research has, therefore, become preoccupied with studies of the contiguity of stimulus and response to the neglect of the structure of relation between words. Examining this verbal edifice reveals that word associations result not from learned stimulus-response pairs, but rather from mental operations of grammar and semantics. These mental operations relate closely to the foundations of our thought processes.

Structure is revealed by considering the similarity of stimulus words, measured here by the "associative meaning" of the words. The associative meaning of any word is the distribution of its associative responses; two words are similar to the extent that their word associations are the same.

This measure of similarity provides for any set of words a complete quantitative description of how much each relates to the others. Particular sets of words are highly interrelated and a factor analysis on such a set produces major components reasonable for the set of words. Applying this technique to a wide variety of words displays certain types of highly related groups of words. These types divide along grammatical lines: the factors for adjectives form scales of polar opposites, e.g. good-bad; prepositions and auxiliary verbs cluster along factors involving grammatical forms, such as active-passive; for nouns and verbs the factors generally follow categories of similarity among the words involved, such as mammals as opposed to birds.

These semantic structures reveal the mental operations which form associations: (1) Contrast, and (2) Grouping, Deese's new laws of association. A word may take its "associative meaning" by contrast provided this is "unique," i.e. the only possible opposite response (for example, pairs of adjectives which are polar opposites); the word may also take its meaning by grouping with other similar words, such as the typical responses to nouns: table-chair, boy-girl, etc.

These laws differ little from the old idea of near and distant responses. They are somewhat more precise than most categories of association, but only because "contrast" is severely restricted by the "uniqueness" requirement. Presumably everything else is grouping. These have an objective basis not usually found in catalogues of associations. This advantage relies upon the component factors of the sets of highly interrelated words. But on what basis are these sets of words chosen? A set of only adjectives could hardly fail to provide factors of contrasting pairs. It is not clear what factors, if any, would result from an arbitrary set of words.

Many of the experiments in the book evidence a disappointing lack of consideration. In particular the sampling techniques are sometimes inadequate. Consider, for example, the argument that word substitution in a sentence frame affects association. A subject constructed sentences using one of an adjective pair. The adjective was removed, and 50 other subjects attempted to guess the missing word. Surely these sentences cannot be considered representative of English usage; when a subject is asked to construct sentences using a particular word, he will most likely attempt to use that word in a fairly distinctive manner. Why has the author avoided the obvious random sampling approach?

This is a good book for those who view association as a process related to categories of response. It provides some interesting comparisons of stimulus words and does so in a quantitative and fairly objective manner. The argument of the book, however, fails; it neither convinces one that associations represent the structure of mental operations nor does it suggest reasons for preferring the structural interpretation to that of stimulus-response theory. It somehow seems that the book both fails to grasp the subtleties of the contiguity theories for association of ideas and also fails to come to grips with the problems of interpreting word-association.

R. P. ALLEN.

A Source Book in the History of Psychology. Edited by R. J. Herrnstein and E. J. Boring. London: Oxford University Press; Harvard University Press. 1965. Pp. xvii + 636. £5.

When Professor Boring published his *History of Experimental Psychology* there was much to relate, and to fit into a framework linking it with times before Fechner's establishment of the formal laboratory discipline. Brett's *History*, concerned with ideas rather than empirically based conclusions, could afford to be less dogged in trying to find a place for everything with everything in its place, and subtler threads could be seen. But

Boring's task was more constrained. Continuity with the philosophical origins of psychological formulation was required. But experimentation, even when it misfires, sharpens conceptual disjunctions uncomfortably. Biological influences, too, raise issues which play havoc with previously accepted rules of the game. Boring was singularly successful in 1929 in producing a book that was clear in standpoint and exposition, full of tempting references to original sources, and eminently readable. Contemporary experimental psychology could be seen as developing naturally out of a "pre-scientific" but rational era through the impact of sixteenth and seventeenth century observation, mathematics and experiment, with a continuing core of philosophical cogitation. This is an orderly and comfortable view—at least it was then to anyone like myself who came to psychology with a background of some physics and a smattering of philosophy. Indeed a great deal, even of more recent development, can be made to fit into it. But inevitably it also excludes or obscures much that is important and more that is interesting. And it has an unfortunate tendency to perpetuate its own pattern.

This book of Readings—for it is open to question whether that is not, by reason of the selection and multiplicity of its contents, a more appropriate title—avowedly aims at a range of cover extending beyond what can be called "experimental." Nevertheless the original outlines of the *History* show through, and the picture which emerges of what William James called "a hope of a science" is informed with that splendidly vigorous and progressive, yet rather restricted, optimism so well-embodied in the Tower named after him.

The book comprises 116 excerpts divided among fifteen chapters. Each excerpt is—obviously—brief, but each makes its point and forms a link in a chain which, though basically simple, does take account of the classical divergencies such as that between the synthetic and the analytic interpretation of mental function. Translation—much of it freshly done—and documentation are well looked after and the introductory and connecting passages by the Editors are helpful and free from undue gloss. What of the selection itself?

Each chapter is self-contained and has its own chronology running from 350 B.C. (where appropriate) to the end of the nineteenth century, with occasional forward excursions into the twentieth to round off a theme. Chapter headings define topics such as "Psycho-physics and Sensory Measurement," "Evolution and Individual Differences," "Functionalism" and "The Reflex" much along the lines of a useful course of introductory lectures. The excerpts themselves have been most skillfully chosen so that none are misleading and few difficult on account of their inevitably meagre context. There are not many intriguing surprises to be found by anyone not quite a newcomer to the history of psychology. But it is pleasing to see some justice done to Hughlings Jackson and Gall. William Small's part in the introduction of the maze into animal learning research is a beautiful example of the pioneering study that is only possible before the well of truth becomes muddied by clumsy stirring. As a sidelight on Titchener, Stella Sharp's pursuit of the individual differences has a certain historical interest. Slightly surprising in the opposite direction is, for instance, the omission in the chapter on cerebral localization of material from Vesalius, or Thomas Willis, while to omit something from Charles Bonnet in the Chapter entitled "Comparative Psychology" seems a pity. But by and large all is much as might be expected having regard to the general framework. Anyone might feel that this or that excerpt could profitably be replaced by some other but, given this framework, no one is likely to cavil generally.

A far more important consideration is that of what form and content a book with this title should have. Or, perhaps more to the point, what form and content *could* it have? How far, for example, should it present material which, though it was ignored at the time and thus played no part in the manifest chain of development, can now be seen to be more aptly relevant to the present state of the subject than much that was in the swim and thus has an accepted place in the traditional story? To what extent is it possible and profitable to introduce those concurrent trends in the scientific and cultural milieu which are often so much more significant than events in the main stream?

What *is* the "Psychology" of today which is supposed to have a "History"? Does it, for example, include features which directly or indirectly have to do with the pathology of mind? Such questions are legion as no one can know better than Professor Boring. And no one, not even he, could produce a book of "Readings" or "Sources" which could invalidate them all. But one wonders whether the time has not come to hope that young and qualified scholars will open up those features of the history of psychology which are at present concealed in the shadows cast by its traditionally accepted worthies. Many

historians of science are, after all, engaged in work drearier and less potentially illuminating than this would be.

Even with its present limitations, however, this book represents an immense amount of erudition and loving care and it has a personal flavour about it which could not have been contrived merely to meet a need. The history of their subject is, in this country at any rate, not a matter for great enthusiasm among psychological students today, who have so much to captivate them in the contemporary scene. But any that look inside this book will be bound to find something to catch his attention and lead him to explore a little. No Departmental Library can afford to be without this book.

R. C. OLDFIELD.

PUBLICATIONS RECEIVED

JOURNALS

International Journal of Psychology, Vol. 1, No. 1, 1966. Edited by G. de Montmollin. Published for the International Union of Psychological Science by Dunod, 92 rue Bonaparte, Paris 6e. Quarterly, one volume per year. Subscription 35F, single issues 9F.

Science Citation Index 1965, *Quarterly No.* 3. *An International Interdisciplinary Index to the Literature of Science and Technology.*
Source Index. Anonymous and A–Zysk.
Citation Index. Aach–Metz.
Citation Index. Metz–Zyus and Anonymous. Patent Citation Index. Corporate Index.
Philadelphia: Institute for Scientific Information Inc.

Mental Health Book Review Index, Vol. 11, 1966. With an editorial: Links between the Humanities and the Literature of the Human Sciences. xxi + 81 pp. Published by the Council on Research in Bibliography, Inc., c/o Research Center for Mental Health, New York University, New York, N.Y. 10003. Sponsored by the World Federation for Mental Health, International Council of Psychologists, and Research Centre for Mental Health, New York University. Price $5.00. Mental Health Book Review Index, c/o Paul Klapper Library, Queens College, Flushing, New York 11367, U.S.A.

Psychology Quarterly, Vol. 2, No. 2, Sept. 1966. Edited by Mohammad Ajmal. Research Journal of the Department of Psychology, Government College, Lahore. Published by Pakistan International Printers, 118, G.T. Road, Lahore.

MONOGRAPHS

The Hampstead Psychoanalytic Index: A Study of the Psychoanalytic Case Material of a Two-Year-Old Child. by John Bolland, Joseph Sandler, *et al.* The Psychoanalytic Study of the Child. Monograph No. 1. New York: International Universities Press. London: Bailey Bros. & Swinfen.) 1965. Pp. xv + 205. $4.50 (40s.).

Pain and Pleasure. By N. R. W. Pande. Monograph read in the Psychology Section of the Indian Philosophical Congress, 1966. New Delhi: Ministry of Defence. Pp. 19.

The Evaluation of the Built Environment. By R. G. Hopkinson. Inaugural Lecture at University College, London. London: H. K. Lewis. 1966. Pp. 27. 5s.

TESTS

A Standardized Road-Map Test of Direction Sense. By John Money. London: Oxford University Press. (Baltimore: Johns Hopkins Press.) 1966. 52s. per set.

Minnesota Test for Differential Diagnosis of Aphasia. By Hildred Schuell. 1966. London: Oxford University Press (Minnesota University press). £5 per set.

BOOKS

Review in a later issue is not precluded by notice here

Working with Groups: Group Process and Individual Growth. (Second Edition). By Walter F. Lifton. London and New York: Wiley. 1966. Pp. xii + 288. 53s.

Recidivism: A Deficiency Disease. By Alastair W. MacLeod. London: Oxford University Press. 1966. Pp. 131. 36s.

Clinical and Social Judgment: The Discrimination of Behavioral Information. By J. Beiri, A. L. Atkins, S. Briar, R. L. Leaman, H. Miller and T. Tripodi. London and New York: Wiley. 1966. Pp. xiv + 271. 60s.

Psychotherapy and the Psychology of Behavior Change. By Arnold P. Goldstein, Kenneth Heller and Lee B. Sechrest. London and New York: Wiley. 1966. Pp. ix + 472. 68s.

A Theory of Achievement Motivation. Edited by John W. Atkinson and Norman T. Feather. London and New York: Wiley. 1966. Pp. x + 392. 87s.

Principles of Psychological Measurement. By G. C. Helmstadter. London: Methuen. 1966. Pp. xx + 248. 36s.

Practical Care of the Mentally Retarded and Mentally Ill. By Ruthanna Penny. Springfield, Illinois: Thomas. 1966. Pp. xviii + 252. $9.50.

Der Sensitive Beziehungswahn: Ein Beitrag zur Paranoiafrage und zur psychiatrischen Charakterlehre. By Ernst Kretschmer. Berlin: Springe-Verlag. 1966. Pp. viii + 236. DM 36.

The Female Alcoholic: A Social Psychological Study. By Barry A. Kinsey. Springfield, Illinois: Thomas. 1966. Pp. viii + 223. $8.75.

Studies on Psychosis: Descriptive, Psycho-Analytic and Psychological Aspects. By Thomas Freeman, John L. Cameron and Andrew McGhie. London: Tavistock. 1965. Pp. vii + 245. 35s.

The First Year of Life: A Psychoanalytic Study of Normal and Deviant Development of Object Relations. By Rene A. Spitz. In collaboration with W. Godfrey Cobliner. New York: International Universities Press. (London: Bailey Bros. & Swinfen.) 1965. Pp. xix + 394. 77s.

Marriage and Family Relations. By Robert C. Williamson. London and New York: Wiley. 1966. Pp. x + 618. 60s.

Socialization after Childhood: Two Essays. By Orville G. Brim, Jr. and Stanton Wheeler. London and New York: Wiley. 1966. Pp. viii + 116. 38s. cloth; 17s. paper.

Group Values through Children's Drawings. By Wayne Dennis. London and New York: Wiley. 1966. Pp. xiii + 211. 53s.

Psychoanalysis in America: Historical Perspectives. Edited by Murray H. Sherman. Springfield, Illinois: Thomas. 1966. Pp. xii + 518. $15.75.

Group Therapy in Childhood Psychosis. By Rex W. Speers and Cornelius Lansing. London: Oxford University Press (North Carolina University Press). 1966. Pp. xx + 186. 38s.

The Excitement and Fascination of Science: A Collection of Autobiographical and Philosophical Essays. By Contemporary Scientists. Palo Alto: Annual Reviews. 1965. Pp. x + 566. Cloth $5.50. Paper $2.45.

The Importance of Lying. By Arnold M. Ludwig. Springfield, Illinois: Thomas. 1965. Pp. xi + 238. $9.75.

Emotions and the Job. By S. G. Rogg and C. A. D'Alonzo. Springfield, Illinois: Thomas. 1965. Pp. xiii + 192. $6.75.

The Wantons. By Lieta Harrison. Trans. by Diana Fussell. London: Ortolan Press. 1966. Pp. 140. 21s.

Absent: School Refusal as an Expression of Disturbed Family Relationships. By Max B. Cline. London: Tavistock Publications. 1966. Pp. xi + 269. 38s.

Leadership and Motivation. By Douglas McGregor. Cambridge, Mass. and London: MIT Press. 1966. Pp. xiii + 286. $6.95, 53s.

Praktikum der Experimentellen Psychologie. By P. Fraisse. Bern and Stuttgart: Hans Huber. 1966. Pp. xv + 379. DM 38.

Psychology in Community Settings: Clinical Educational, Vocational, Social Aspects. By Seymour B. Sarason, Murray Levine, I. Ira Goldenberg, Dennis L. Cherlin and Edward M. Bennett. London and New York: Wiley. 1966. Pp. x + 714. 98s.

Theories in Contemporary Psychology. Edited by Melvin H. Marx. London: Collier-Macmillan. 1963. Reprinted 1965. Pp. xi + 428. 80s.

Recent Advances in Biological Psychiatry. Vol. VIII. Edited by Joseph Wortis. Proceedings of the 20th Annual Convention and Scientific Program of the Society of Biological Psychiatry. New York: Plenum Publishing Corp. 1966. Pp. xiii + 367. $16.50.

Human Spatial Orientation. By I. P. Howard and W. B. Templeton. London and New York: 1966. Pp. 533. 84s.

Foundations of Clinical Psychology. By Richard H. Dana. London: Van Nostrand. 1966. Pp. viii + 322. 54s.

The Genesis of Language: A Psycholinguistic Approach. Edited by Frank Smith and George A. Miller. Cambridge, Mass and London: M.I.T. Press. 1966. Pp. xii + 400. 80s.

PAPERBACKS

The Fields of Group Psychotherapy. Edited by S. R. Slavson. London and New York: Wiley. Science Editions. 1966. Pp. xiii + 338. 15s. ($1.95).

Research Techniques in Human Engineering. By Alphonse Chapanis. London: Oxford University Press (Johns Hopkins Press). 1966. Pp. xii + 316. 20s. ($2.45). *Reviewed in this Journal, Vol. XII, 1960, p. 192.*

Check Your Own I.Q. By H. J. Eysenck. Harmondsworth: Penguin Books. Pelican Original. 1966. Pp. 192. 3s. 6d.

Some Problems of Measurement in Psychophysics. By Kenneth Junge. Oslo: Universitetsforlaget. 1966. Pp. 61. n.kr. 18.0.

The Biology of Art. By Desmond Morris. London: Methuen. University Paperbacks. 1966. Pp. 176. 25s. (cloth, 42s.). *Reviewed in this Journal, Vol. XV, 1963, p. 80.*

Anxiety, Nervousness and Depression. By F. R. C. Casson. London: British Medical Association. Family Doctor Booklet. 1966. Pp. 31. 1s.

A Programmed Introduction to Statistics. By Freeman F. Elzey. Belmont, California: Wadsworth. 1966. Pp. viii + 376.

Hypnosis and Related States: Psychoanalytic Studies in Regression. By Merton M. Gill and Margaret Brenman. London and New York: Wiley. Science Editions. 1966. Pp. xxiv + 404. 17s. ($2.25).

Inside the Black Room: Studies of Sensory Deprivation. By Jack Vernon. London Penguin Books. Pelican. 1966. Pp. 182. 3s. 6d. *Reviewed in this Journal, Vol. XVII, p. 94.*

Psychologia a Kybernetika. By Michal Strizenec. Bratislava: Slovak Publishing House. 1966. Pp. 241. Kcs 23. (10-page English summary.)

Kompendium der psychodiagnostischen Tests 2. neubearbeitete und erweiterte Auflage. By H. Hiltmann. Bern and Stuttgard: Bern. 1966. Pp. 304. Fr./DM 22.00.

Insight Books. London: Van Nostrand. 1966.

No. 32. *Interpersonal Accommodation.* By Mark Abrahamson. Pp. vii + 151. 14s. $1.75.

No. 33. *Ward H: An Adventure in Innovation.* By N. J. Colarelli and Saul M. Siegel. Pp. viii + 248. 18s. $2.25.

No. 34. *Attention.* Edited by Paul Bakan. Pp. vi + 225. 14s. $1.75.

No. 35. *Reconciliation: A Theory of Man Transcending.* Edited by Sidney M. Jourard and Dan C. Overlade. Pp. x + 148. 14s. $1.75.

Walden Two. By B. F. Skinner. London: Collier-Macmillan. Paperback Edition. 1962. Pp. 320. 16s.

Science and Human Behavior. London: Collier Macmillan. Paperback Edition. 1965. Pp. x + 461. 25s.

Animal Behavior. By Keller Breland and Marian Breland. London: Collier-Macmillan. 1966. Pp. xii + 210. 16s.

Social Psychology: An Experimental Approach. By Robert B. Zajonc. Belmont, California: Wadsworth (London: Prentice-Hall International.) 1966. Pp. viii + 120. 12s. 6d.

PROCEEDINGS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY, 1966

28th–30th March, 1966. Meeting at Bristol.

1st Session: "Evaluation of negatives and passives in relation to meaningful change," by Mrs. Judith Greene.* "The effect of exogenous hormones on the aggressive and defensive behaviour in the Ring Dove," by Dr. D. M. Vowles and Mr. D. Harwood.* "Intracranial self-stimulation behaviour in chilled hibernators," by Dr. N. Mrosovsky.

2nd Session: "Tolerance and cross-tolerance studies between mescaline and its analogues in the rat," by Dr. J. R. Smythies and Miss E. A. Sykes.* "The 'efficiency' of the conditioned eyelid response," by Dr. Irene Martin and Mr. A. Levey.*

3rd Session: "Effect of noise on perceptual selection," by Dr. D. E. Broadbent and Dr. Margaret Gregory.* "Errors and error-correction in choice responses," by Dr. P. M. A. Rabbitt. "Subjects' strategies in mediated facilitation of paired-associate learning," by Dr. R. O. Rouse.*

4th Session: "Performance under progressive interval schedule with and without reset contingency," by Mr. P. Harzem.* "Eye movements in form perception," by Dr. J. A. M. Howe.

5th Session: "Discrimination learning in normal and severely sub-normal children," by Dr. P. Bryant.* "Some effects of bilateral lesions within the frontal or temporal lobes upon the ability of baboons to match and discriminate visual stimuli," by Dr. A. W. H. Buffery.*

11th–13th July, 1966. Meeting at Cambridge.

1st Session: "The interaction of long and short-term memory," by Dr. A. Baddeley. "The manipulation of split-span memory behaviour by pay-off strategies," by Dr. N. Moray. "Recognition of multi-dimensional stimuli," by Dr. D. Corcoran.

2nd Session: "Observing responses and the reduction of uncertainty by the monkey," by Dr. J. Steiner.* "Disturbances of size constancy following cortical lesions in the monkey," by Mr. N. Humphrey.* "A comparison of the effects of striate cortex and retinal lesions on visual acuity in the monkey," by Dr. L. Weiskrantz and Dr. A. Cowey.

3rd Session: "Communication between quail embryos and its effect on hatching time," by Miss M. Vince. "Cardiovascular conditioning," by Professor S. I. Cohen.*
The Sir Frederic Bartlett Lecture "Things, words and the brain," by Professor R. C. Oldfield.

4th Session: "The suppression of visual after-images by eye movements," by Mr. J. H. Garton,* Mr. A. L. Holden* and Mr. D. P. M. Northmore.* "Characteristics of light onset reinforcement," by Mr. C. C. Kiernan.* "Timing behaviour in the rat," by Mr. R. J. Bradley* and Dr. J. R. Smythies. "A unique behavioural profile for psychometric drugs," by Mr. V. S. Johnson* and Dr. J. R. Smythies.

5th Session: "The new Pseudophysics: some models for magnitude estimation," by Dr. E. C. Poulton. "Brightness contrast," by Dr. P. Whittle.

2nd–4th January, 1967. Annual General Meeting at University College London.

1st Session: "Selective inhibition of operant responses by chorpromazine: studies in the albino rat," by Dr. A. R. King.* "Acquisition and extinction response profiles in within-subject experiments involving discontinuous negatively correlated reinforcement to one stimulus and continuous reinforcement to another," by Dr. A. Amsel.* "Texture preferences and attachments of domestic chicks," by Miss Ann Taylor* and Professor W. Sluckin.

2nd Session: "Immediate memory for words and pictures in different groups of children," by Dr. Beate Hermelin and Dr. N. O'Connor. "Visual-proprioceptive correspondence," by Dr. D. Legge. "Human performance in low signal probability tasks," by Dr. P. Lucas.*

3rd Session: "An analysis of serial betting behaviour," by Dr. K. D. McRae.* "Sequential effects in two-choice, forced choice psychophysical procedure," by Dr. I. M. Hughes. "Speech interfering aspects of noises," by Dr. J. C. Webster.* "Stimulus repetition effects and identification of sets of signals," by Dr. P. M. A. Rabbitt.

4th Session: "The Brook Reaction as a test of temperament," by Dr. A. W. Heim. "The role of selective perception in movement and movement after-effects," by Dr. N. F. Dixon.

5th Session: "Exploration of a model of word storage," by Mr. G. R. Kiss.* "Time series analysis of the EEG," by Mrs. Patricia Michie.*

* By invitation.

Committee, 1967: President: Professor J. Brown.

Editor: Mr. A. J. Watson.

Ordinary Members of the Committee: Dr. R. Davis (Hon. Treasurer), Dr. B. M. Hermelin, Professor I. M. L. Hunter, Dr. D. Legge (Asst. Hon. Secretary), Dr. P. M. A. Rabbitt, Professor R. L. Reid, Professor J. R. Symons, Dr. P. H. Venables (Hon. Secretary), Dr. R. T. Wilkinson.

The following have been elected into membership of the Society:

M. Argyle, M.A., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
C. B. Blakemore, B.Sc., Department of Psychology, Institute of Psychiatry, Maudsley Hospital, Denmark Hill, London, S.E.5.
A. W. H. Buffery, M.A., Ph.D., Unit for Research on Medical Applications of Psychology, 5 Salisbury Villas, Station Road, Cambridge.
B. Delisle Burns, B.A., M.R.C.S., L.R.C.P., Division of Physiology and Pharmacology, M.R.C. National Institute for Medical Research, Mill Hill.
P. Bryant, B.A., Ph.D., M.R.C., Social Psychiatry Research Unit, Institute of Psychiatry, Maudsley Hospital, Denmark Hill, London, S.E.5.
J. J. Cowley, B.S.c., M.Sc., Department of Psychology, University of Aberdeen, Aberdeen.
F. I. M. Craik, B.Sc., Ph.D., Department of Psychology, Birkbeck College, London, W.C.1.
K. Gibbins, B.A., Ph.D., Department of Psychology, University of Newcastle upon Tyne, Sydenham Terrace, Newcastle upon Tyne, 2.
M. S. Halliday, B.A., Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
J. A. M. Howe, M.A., Ph.D., Department of Machine Intelligence and Perception, The University, Edinburgh.
Susan D. Iversen, M.A., Ph.D., Psychological Laboratory, Downing Street, Cambridge.
Daphné Joyce, B.Sc., Ph.D., Department of Psychology, Birkbeck College, London, W.C.1.
C. C. Kiernan, B.A., Department of Psychology, Birkbeck College, London, W.C.1.
D. J. Murray, M.A., Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.
Helen E. Ross, B.A., Ph.D., Department of Psychology, The University, Hull.
T. Shallice, B.A., Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

Visiting Foreign Members:

Professor R. Albino, University College, Gower Street, London, W.C.1.
Professor A. Amsel, University College, Gower Street, London, W.C.1.
Professor W. R. Garner, Applied Psychology Research Unit, Cambridge.
Dr. E. P. Lovejoy, Laboratory of Experimental Psychology, Sussex.
Professor T. Reese, Sub-Department of Animal Behaviour, Madingley, Cambridge.
Dr. I. K. Taylor, Department of Communications, Keele.
Dr. M. M. Taylor, Department of Communications, Keele.
Dr. J. C. Webster, Applied Psychology Research Unit, Cambridge.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIX 1967 Part II

REHEARSAL GROUPING AND HIERARCHICAL ORGANIZATION OF SERIAL POSITION CUES IN SHORT-TERM MEMORY

BY

WAYNE A. WICKELGREN

From the Department of Psychology, Massachusetts Institute of Technology

Lists of 8, 9, or 10 digits were presented at the rate of 1 digit/sec. to subjects instructed to rehearse silently the digits in non-overlapping groups of 1, 2, 3, 4, or 5 digits, after hearing each digit. Subjects were instructed not to rehearse any digits outside the group currently being presented. Rehearsing in 3's was optimal, irrespective of list length. Both recall of items and recall of the correct positions of items improved from 1's to 2's to 3's. Recall of items declined very little from 3's to 4's to 5's, but recall of position declined sharply. Errors in positioning digits tended, above chance, to be in the same group or the same position in a different group. The results suggest that both item-to-item associations and serial position-to-item associations are formed in short-term memory, that only two or three serial position cues are used, but that these serial position cues can be hierarchically organized into a beginning, middle, and end group and a beginning, middle, and end position within a group.

INTRODUCTION

What is learned by a subject who is presented with a list of items depends, in the final analysis, upon what he thinks about during presentation of the list, that is, upon what internal representatives are activated in what sequence. If the items in the list are highly intelligible, the list not too long, the subject co-operative, and instructed to attend to each item, it is reasonable to suppose that the internal representative of each item of the list is activated at about the time it is presented. If the rate is fast enough (2 to 4 syllables a sec. or faster), this is very likely to be close to all that is activated, and in such cases the analysis of learning is vastly simplified (e.g. Wickelgren and Norman, 1966; Wickelgren, 1967), though we are still far from a complete understanding of acquisition, storage, and retrieval even in these "simple" situations.

However, when the list is presented at slower rates (1 syllable a sec. or slower), it is introspectively clear that internal representatives are activated which do not correspond to items in the list (partly by association to the internal representatives of one or more items on the list) and, furthermore, the internal representatives of previous items on the list are activated *after* one or more subsequent items have been presented (i.e. more than the present item is rehearsed). Under these circumstances, the analysis of what is learned (what associations are strengthened) is a very difficult problem.

One way to attack the problem is to attempt to control, by instruction, the active rehearsal process. If the instructional manipulation is effective, there should be differences in the rates and types of errors among groups instructed to rehearse in

different ways. Furthermore, the exact nature of these performance differences should indicate some characteristics of man's active rehearsal and recoding capacities in list-learning tasks.

The present study is directed to one dimension of rehearsal strategy, namely, the size of the groups in which the subject rehearses a list. The concern is not with determining what grouping methods subjects will choose, but rather with the consequences for performance of adopting different grouping methods. The study is restricted to the rehearsal of non-overlapping groups of different size. This means that a subject is instructed to rehearse, after each item is presented, all previously presented items of the group to which the present item belongs. Furthermore, he is explicitly instructed not to rehearse items from any other group. Thus, if a subject is rehearsing in 4's, after hearing the eighth digit of a list of 10 digits, he is to rehearse the fifth through eighth items of the list in order and no other items. When the ninth item is presented, he is to rehearse only that item (because now he is starting on the third group).

Under ordinary circumstances when one says that he is grouping in 4's, he does not mean that he follows this rigid method of rehearsal. He means that he "thinks" of the list in groups of four items each (four, four, and two in the case of a list of 10 items). Under ordinary circumstances subjects rehearsing in 4's would not refrain from rehearsing a previous group while being presented with the second group, nor is there any reason to think this would be desirable for memory performance. However, it is very desirable for ease of theoretical analysis to control as carefully as possible the exact method of rehearsal, and this is the reason for doing it. Instructing a subject to rehearse in non-overlapping groups of four items induces a subject to "think" of the list in groups of four, whatever that might mean, and also provides explicit instruction regarding what is to be rehearsed at each moment. The hope is that the analysis of performance under different rehearsal instructions will suggest what it means to "think" of a list in groups of four.

The present study is a replication, under slightly different conditions, of a previous study (Wickelgren, 1964) which demonstrated the following: (*a*) the probability of correct ordered recall is greater for grouping in 3's than for grouping in 1's, 2's, 4's, or 5's; (*b*) while the increase in correct performance from 1's or 2's to 3's is in both recall of items and recall of the correct position of the item in the list, the decrease from 3's to 4's and 5's is primarily in recall of the correct position of items, not the recall of items; and (*c*) errors in positioning items tend, beyond chance, to be in the same group or in the same position of a different group. The results were interpreted as indicating that subjects form associations from serial position cues (such as "beginning," "middle," and "end") to items and that these serial position cues can be hierarchically organized into a beginning, middle and end group and a beginning, middle and end position within a group.

In the previous experiment, each subject was run in only two different rehearsal conditions (first half of the session in one condition, second half in the other), for fear that subjects could not rapidly switch from one rehearsal strategy to another. In the present study, each subject was run under all rehearsal instructions, with the rehearsal method changing every five trials. To facilitate the shift, the boxes in which a subject wrote his answers were grouped in correspondence with the rehearsal instruction. For example:

In the previous experiment, subjects were run on a small number of trials with all list lengths from 6 to 10 digits. It was not possible to analyse the results separately

for each list length. In the present experiment, separate experiments were done using list lengths of 8, 9, and 10 digits to determine whether the conclusions of the previous study held for each list length or whether there was an interaction between rehearsal group size and list length. For example, it seems "natural" to group a list of eight items into two groups of four, a list of nine items into three groups of three, etc.

Method

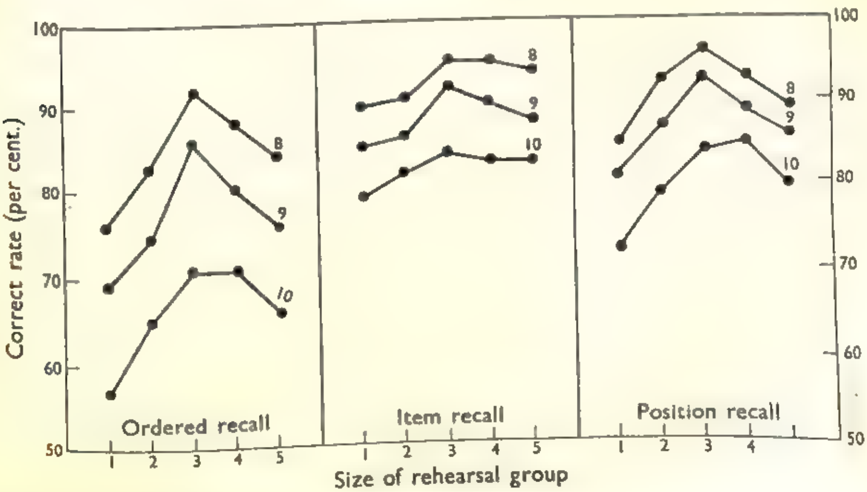
On each trial a subject got a ready signal, followed in about half a sec. by a list of 8, 9, or 10 digits presented at the rate of 1 digit/sec., followed by 10 sec. in which to recall the list in order by filling in boxes on an answer sheet. There were three experiments, one using lists of 8 digits, one using lists of 9 digits, and one using lists of 10 digits. There were no digits repeated in any list, so the lists of 10 are essentially permutations of the 10 digits, 0–9. Within each experiment, subjects used each of the five different grouping methods (rehearsing in 1's, 2's, 3's, 4's, or 5's), changing the grouping method every five trials. In every 25 trials subjects used each of the grouping methods once for five trials in a block, with the order of the grouping methods randomized. Furthermore, the subjects in each experiment were divided roughly in half and one half had the order of the grouping methods exactly reversed from the other half.

Subjects were M.I.T. undergraduates taking psychology courses, who participated in the experiment as part of their course requirements; 23 took the length-8 lists, 24 the length-9 lists, and 30 the length-10 lists.

Results

The data were analysed for ordered recall, item recall, and position recall of individual items in the list. An item is correct by an ordered recall criterion if it appears in the correct box on the answer sheet. An item is correct by an item recall criterion if it appears in any box on the answer sheet. The correct position recall rate is obtained by dividing the number of items correct by an ordered-recall criterion by the number of items correct by an item-recall criterion. Thus, item recall and position recall are statistically independent, and ordered recall combines these two factors into an overall score. This analysis was done separately for each

FIGURE 1



Effect of size of rehearsal group on ordered recall, item recall, and position recall, for lists of 8, 9, and 10 digits.

subject for each rehearsal grouping method to permit comparison of performance under different grouping methods by the Wilcoxon Matched-Pairs Signed-Ranks test.

The average probabilities of correct ordered, item, and position recall for each rehearsal grouping method for each list length are shown in Figure 1. The results agree in almost every detail with the previous study (Wickelgren, 1964). Furthermore, the efficacy of different grouping methods is remarkably invariant over different list lengths. Rehearsing in 3's is optimal regardless of list length, although it is no better than 4's for length-10 lists. Accuracy in both item and position recall increases from 1's to 2's to 3's for all three list lengths. Accuracy in position recall then declines from 3's to 4's to 5's for all list lengths, with one small reversal at length-10. However, item recall remains roughly constant from 3's to 4's to 5's for lengths 8 and 10, and declines slightly for length 9.

By the Wilcoxon test, R1 < R2 in ordered recall ($p < 0{\cdot}01$ for all three list lengths), in item recall (correct direction but insignificant for all three lengths), and in position recall ($p < 0{\cdot}01$ for all three lengths). R2 < R3 in ordered recall ($p < 0{\cdot}01$ for all three lengths), in item recall ($p < 0{\cdot}01$ for all three list lengths), and in position recall ($p < 0{\cdot}01$, 0·01, and 0·05 for lengths 8, 9, and 10, respectively). R3 > R4 in ordered recall (0·05, 0·01, and n.s.), item recall (n.s., 0·05, n.s.), and position recall (0·01, 0·01, n.s.). R4 > R5 in ordered recall (0·01, 0·05, 0·05), item recall (n.s., n.s., n.s.), and position recall (0·01, 0·01, 0·05).

To obtain more information about the curvilinear relationship between size of rehearsal group and position recall, every item that was recalled correctly by an item-recall criterion, but recalled in the wrong position, was classified into one of three position-error categories. *Within-group* errors refer to items recalled in the correct group, but at the wrong position within the group. *Within-position* errors refer to items recalled in the wrong group, but in the correct position within the group. *Other* errors refer to items recalled in the wrong group and wrong position within the group.

To obtain an independent measure of the relative frequency of within-group and within-position errors for Conditions R2, R3, R4, and R5, the frequency of each type of error was compared to the frequency of all *other* errors. In order to determine if the relative frequency of within-group or within-position errors was affected by rehearsal group size, it was necessary to compare the relative frequency of each type of error in any rehearsal condition to the same relative frequency computed for a standard "ungrouped" condition, in this case R1. Computation of the comparable relative frequency in R1 involves breaking R1 into the same groups as in the condition with which it is being compared.

Table I presents the relative frequency of within-group and within-position errors in each condition for each list length and the comparable relative frequency in R1.

For all list lengths and for all rehearsal conditions, there is clearly a tendency for errors to be in the same group or in the same position within a group. As in the previous study, the tendency to make within-group errors increases with the number of positions in a group and the tendency to make within-position errors increases with the number of groups.

The findings of the previous study are completely replicated. Subjects appear to have little difficulty in rapidly changing rehearsal set. Furthermore, there appears to be no significance for memory performance of the greater "naturalness" of dividing a list of eight items into two groups of four or of dividing a list of 10 items into two groups of five.

TABLE I
Errors in Positioning Items

Condition	List length	Within-Group/Other: Error ratios in R2 to R5	Within-Group/Other: Comparable error ratio in R1	Within-Position/Other: Error ratios in R2 to R5	Within-Position/Other: Comparable error ratio in R1
R2	8	0·66*	0·47	1·19**	0·53
	9	0·49	0·49	1·55**	0·65
	10	0·51	0·46	2·11**	0·81
R3	8	0·71	0·80	0·62	0·22
	9	1·18**	0·93	1·04**	0·37
	10	1·30*	0·88	1·28**	0·39
R4	8	2·19	1·58	0·21	0·12
	9	1·99**	1·40	0·33**	0·20
	10	1·57*	1·11	0·33**	0·19
R5	8	2·99**	1·54	0·12*	0·03
	9	2·02*	1·74	0·19**	0·05
	10	2·96**	1·96	0·25*	0·15

Note. Error ratios in R2 to R5 that are significantly greater than the comparable error ratios in R1 are indicated by asterisks (Wilcoxon Matched-Pairs Signed-Ranks test).

* $p < 0{\cdot}05$.
** $p < 0{\cdot}01$.

Discussion

If item-to-item associations were the only memory traces mediating performance in immediate memory tasks, one would expect performance to improve with increasing size of rehearsal group. This is because the ratio of the number of strengthened item-to-item associations to the number of rehearsed items becomes progressively more favourable with increasing size of rehearsal group. Furthermore, there is no reason to expect significant numbers of within-position errors, if item-to-item associations are the only associations mediating immediate memory performance.

Thus, there is reason to believe that serial position-to-item associations are also contributing to correct memory performance in this situation. In addition, it is necessary to assume that these serial position cues can be hierarchically organized into a "beginning," "middle," and "end" group and a "beginning," "middle," and "end" position within a group. Such a theory gives a satisfactory intuitive explanation of why there is an optimum size of rehearsal group. The explanation is as follows: when either the number of positions within a group or the number of groups gets too high, the mapping from serial position cues to items becomes too ambiguous, assuming that humans possess only a finite number of serial position concepts.

Because the optimum rehearsal group size was three in the present experiments, the indicated number of different serial position concepts is in the neighbourhood of three. Because of the possibility of defining one serial position concept (e.g. "middle") by exclusion, the possibility of using remote serial position-to-item associations, and the availability of item-to-item associations, it is at least equally likely that there are only two serial position concepts used by humans in this situation ("beginning" and "end"), as that there are three serial position concepts. A more precise formulation of the theory is needed to settle such a question.

The present results should not be interpreted as indicating that serial position cues are the only cues mediating short-term serial list-learning. The "associative intrusions" found in other studies (Wickelgren, 1965, and 1966) argue strongly for the existence of item-to-item associations. Furthermore, the relatively negligible decline in item recall from 3's to 4's and 5's suggests the existence of a compensating factor which is improving with increasing group size, namely, item-to-item associations. Available evidence indicates that both serial position-to-item associations and item-to-item associations are strengthened by presentation of a list and that both are used in retrieval from short-term memory.

Finally, it is worth noting that the associative-memory conception of serial-position information is completely compatible with the simultaneous use of item-to-item information. Non-associative memories, in which serial position is easily represented by an ordered set of locations, are not very easily adapted to make simultaneous use of item-to-item information.

This research was supported primarily by grant, MH 08890–03, from the National Institute of Mental Health, Public Health Service. Further aid was received from a National Aeronautics and Space Administration grant, NsG 496, to Hans-Lukas Teuber.

References

Wickelgren, W. A. (1964). Size of rehearsal group and short-term memory. *J. exp. Psychol.*, **68**, 413–9.

Wickelgren, W. A. (1965). Short-term memory for repeated and non-repeated items. *Quart. J. exp. Psychol.*, **17**, 14–25.

Wickelgren, W. A., and Norman, D. A. (1966). Strength models and serial position in short-term recognition memory. *J. Math. Psychol.*, **3**, 316–47.

Wickelgren, W. A. (1966). Associative intrusions in short-term recall. *J. exp. Psychol.*, **72**, 853–8.

Wickelgren, W. A. (1967). Exponential decay and independence from irrelevant associations in short-term recognition memory for serial order. *J. exp. Psychol.*, **73**, 165–71.

Manuscript received 24th June, 1966.

FAMILIARITY AND FREE RECALL

BY

H. C. A. DALE*

From the Medical Research Council's Applied Psychology Research Unit, Cambridge

Using stimulus material drawn from a single category, it is shown that formal recognition tests yield superior performance to recall when the category cannot be completely enumerated by the subjects, but not otherwise. Furthermore, within these categories recognition is superior only if unfamiliar items are selected as stimuli. These findings fit the theory that when recalling subjects attempt to scan the relevant category and "recognize" those items employed as stimuli.

Introduction

As Tulving (1964) has noted, in most laboratory free recall learning experiments the stimuli are not being learned, since they are already known; instead, the subjects are merely required to link them with a specific experimental situation. In these experiments performance is positively correlated with the degree of prior familiarity of the stimuli (cf. Poulton, 1957; Underwood and Schulz, 1960). Particularly relevant to the present investigation is the observation by the writer that this is true when all stimuli are drawn from a single category. In a recall test, items which were frequently given by similar subjects when required to list the complete category were commonly recalled correctly, whereas items rarely listed were less readily remembered (Dale, 1967).

The present study examines the possibility that this facilitatory effect of familiarity arises during retrieval. It is hypothesized that when retrieving the subject attempts to scan all members of the category or categories from which the stimuli have been drawn. He then endeavours to identify the stimuli. This is equivalent to asserting the subject converts the recall test into a recognition test in which he generates his own distractors.

According to this view, successful retrieval of a particular item can only occur when two conditions are satisfied:

(*a*) the item is included amongst those scanned, and

(*b*) it is successfully identified when scanned.

Familiarity could possibly affect the probability that an item is amongst those available for scanning and in this way influence performance. If this is so, the effect of familiarity could be eliminated by using a recognition test in which the complete set of category members is used as the array of recognition alternatives. A number of predictions can be made if familiarity contributes in this way: (1) the degree of familiarity of a stimulus would not be correlated with its probability of successful recognition; (2) recognition performance would be superior to recall with unfamiliar stimuli but not with familiar stimuli (since the latter would be scanned in recall); (3) with categories in which all items are familiar, recognition would be no better than recall.

Davis, Sutherland and Judd (1961) have reported an experiment which lends support to the third prediction. Using a list of 15 two-digit numbers as stimuli,

* Now at Department of Psychology, University of Hull.

they found that with an immediate test, raw free recall scores were as high as raw recognition scores with a complete-category recognition test. This result could possibly have depended upon the particular conditions employed. The present investigation began, therefore, by repeating this experiment under a wider range of conditions.

Experiment I

Method

Procedure. All conditions consisted of a single presentation of a list of stimuli and a single retention test. Retention was measured either by free recall or by a complete category recognition test. Subjects were tested in groups of from 8 to 30 and groups were always split, alternate subjects being given recognition tests while the remainder were tested by free recall. The other independent variables were as follows: (1) time of test, immediate or after 40 min. spent on other unrelated experimental tasks; (2) presentation rate, 60 per min. or 10 per min.; (3) mode of presentation, visual or auditory. With visual presentation the stimuli were printed on large (10 in. × 24 in.) cards from 2·5 in. stencils. When presented rapidly the cards were stacked and flipped over. When presented slowly each was held up by the experimenter for 5 sec. with a 1 sec. interval between successive stimuli. With the auditory presentation each stimulus was read aloud once by the experimenter. The two-digit numbers used as stimuli were presented in two different orders. Order 1 was: 68, 42, 11, 83, 95, 26, 58, 97, 59, 32, 36, 13, 28, 33, 51. Order 2 was: 95, 58, 97, 59, 68, 42, 51, 26, 36, 13, 28, 33, 11, 32, 83. In the recognition tests all possible two-digit numbers were printed in ascending order in columns of 10. In order to conform with the technique employed by Davis *et al.* the printed sheets in the experiment carried the instruction: "If you are not sure, guess." All subjects were reminded during the 3 min. allowed for the retention test that 15 two-digit numbers had been presented.

Forty fresh subjects were used with each condition, 20 having a recall test and 20 a recognition test. All subjects were young, newly-enlisted, men.

Results and discussion

Mean raw scores together with the mean total number of responses are tabulated in Table I. The main results of an analysis of variance on an arc-sine transformation

TABLE I

Memory for 15 Two-Digit Numbers Showing Mean Number Correctly Recalled or Recognized (S) and the Mean Total Number of Responses (T) (20 subjects per observation)

Time of test	*Rate of presentation*	*Mode of presentation*	*Order 1*				*Order 2*			
			Recall		*Recognition*		*Recall*		*Recognition*	
			S	*T*	*S*	*T*	*S*	*T*	*S*	*T*
Immediate ..	Rapid (60/min.)	Visual	6·85	11·05	7·10	14·05	7·10	11·25	7·15	13·45
		Auditory	5·55	10·95	5·02	14·15	6·30	11·80	6·10	12·50
	Slow (10/min.)	Visual	9·85	13·40	9·00	13·65	9·00	13·00	8·75	14·30
		Auditory	8·65	14·05	7·85	14·30	8·75	12·15	9·20	14·05
After 40 min. delay ..	Rapid (60/min.)	Visual	6·10	11·30	6·00	13·60	6·05	11·80	6·55	13·50
		Auditory	4·60	12·00	6·00	14·30	5·00	10·35	4·70	13·35
	Slow (10/min.)	Visual	7·15	12·60	8·10	13·80	8·60	13·65	8·60	14·80
		Auditory	7·65	12·40	6·70	14·80	8·55	13·25	8·60	14·65

of the scores are given in Table IV. In addition to the main effects, the following interactions were significant: R × P ($p < 0·025$), and T × R × O ($p < 0·025$). It must be stressed that no advantage of recognition was obtained despite the fact that raw scores were used as a measure. More responses were made with the recognition test procedure than in recall, but these did not lead to higher scores.

This experiment has shown that for digits Davis *et al.* result holds under a variety

of conditions. Thus it has been shown that when a category is completely known a recognition test offers no advantage over recall.

The second experiment to be reported examines the other two predictions, namely that when the stimuli are drawn from a category in which the members are not all highly familiar: (1) differences in familiarity will affect the probability of recall, but not that of recognition; and (2) recognition will be superior to recall if unfamiliar stimuli are used, but not otherwise.

Experiment II

In this experiment the recall-recognition comparison was made for a sample of English County names. Although the English Counties would generally be known by the subjects, it is unlikely that any subject could enumerate the category completely. To check this supposition a preliminary test was carried out on a separate group of 86 subjects who were not used in the subsequent memory tests. All subjects were drawn from the same source of young newly-enlisted men.

Method

Assessment of background knowledge. Eighty-six subjects were requested to list all the English Counties. They were tested in groups of from 10 to 30 and were unaware both of the time which would be allowed (10 min.) and of the total category size (40). The mean number of Counties listed by each subject was 19·99, S.D. = 5·45. The number of subjects listing each category member is given in Table II.

TABLE II

THE FREQUENCY WITH WHICH EACH COUNTY WAS LISTED

*Hampshire	80	**Gloucestershire	39
*Yorkshire	79	**Leicestershire	37
Berkshire	76	Northamptonshire	36
**Lancashire	75	Rutland	36
*Devonshire	72	**Cambridgeshire	35
*Essex	68	**Cheshire	35
Sussex	68	**Cumberland	34
*Cornwall	64	Suffolk	34
*Kent	61	Dorset	32
**Derbyshire	58	Wiltshire	30
*Middlesex	55	Worcestershire	30
*Surrey	54	Buckinghamshire	26
**Lincolnshire	53	*Shropshire	26
**Nottinghamshire	52	**Staffordshire	26
**Somerset	48	*Hertfordshire	25
Warwickshire	47	Bedfordshire	21
Northumberland	43	Herefordshire	17
*Norfolk	39	*Westmorland	15
Oxfordshire	39	Huntingdonshire	8
**Durham	39	Monmouthshire	7

* Used in List 1.
** Used in List 2.

Memory tests. Two samples of 12 County names were used as stimuli. They are marked with asterisks in Table II. With each list, two orders of presentation were employed, one being the reverse of the other. On the recognition test paper all 40 Counties were printed in alphabetical order. The sheet did not carry a guessing instruction. The general conditions paralleled those of Experiment I. During the 3 min. retention

test the subjects were reminded that 12 English County names had been presented. 640 fresh subjects were employed.

Results

Raw scores are given in Table III together with values of t for the significance of the regression of performance upon familiarity. The regression coefficients were generally positive for recall and zero for recognition. The results of the analysis of variance show that this difference is highly significant (cf. Table IV).

TABLE III

MEMORY FOR 12 ENGLISH COUNTY NAMES, SHOWING MEAN RAW RECALL AND RECOGNITION SCORES, MEAN NUMBER OF RESPONSES AND t, WHERE $t = \beta/\text{S.E.}$, β BEING THE REGRESSION OF PERFORMANCE UPON RESPONSE AVAILABILITY WITH LIST POSITION EFFECTS CONTROLLED

(Each mean is based upon 20 subjects' responses)

Time of test	*Rate of presentation*	*Mode of presentation*	*List 1*						*List 2*					
			Recall			*Recognition*			*Recall*			*Recognition*		
			Score	*Total*	t*	*Score*	*Total*	t	*Score*	*Total*	t	*Score*	*Total*	t
Immediate	Rapid (60/min.)	Visual	8·40	9·30	+0·80	7·85	10·95	−0·50	6·45	7·70	+1·16	9·05	11·60	−0·67
		Auditory	7·75	9·85	+0·33	8·55	11·75	+1·01	7·65	9·75	+1·10	9·05	11·60	−0·33
	Slow (10/min.)	Visual	9·60	10·10	−0·71	10·20	11·50	−0·04	7·95	8·50	+1·32	10·45	11·85	−2·71
		Auditory	9·85	10·60	+1·23	9·40	11·85	+1·39	8·35	8·95	+1·05	10·75	11·65	+0·58
After 40 min. delay	Rapid (60/min.)	Visual	7·00	8·25	+1·80	7·80	11·30	−0·08	6·05	8·10	+1·43	8·75	12·05	+2·30
		Auditory	7·15	9·30	+4·22	7·75	11·85	+0·22	6·40	7·95	+1·91	8·20	11·95	−0·85
	Slow (10/min.)	Visual	8·10	9·05	+1·81	9·65	12·00	−0·82	8·55	9·55	+1·32	10·40	11·75	−1·38
		Auditory	9·00	9·80	+3·38	9·90	12·00	+2·13	8·85	9·85	+2·21	10·15	11·90	−0·54
Overall mean	..	..	8·36	9·53	—	8·89	11·65	—	7·53	8·79	—	9·60	11·79	—

* Where $t > 2{\cdot}09$, $p < 0{\cdot}05$; $t > 2{\cdot}85$, $p < 0{\cdot}01$.

The major finding of the between-conditions analysis is that in general recognition yields higher scores than recall. A summary of the main effects together with significance levels determined from analyses of variance on transformed data is given in Table IV. In addition to these main effects the following interactions were significant: M × O ($p < 0{\cdot}001$), and T × M × O ($p < 0{\cdot}01$). The M × O interaction indicates that the superiority of recognition was greater with List 2 The T × M × O interaction arises because the M × O interaction is more marked on the immediate test than after a delay.

On immediate test, the advantage of recognition showed only with List 2, which consisted of relatively unfamiliar stimuli. The extent to which the Counties in List 1 were generally better known than those in List 2 was not great. The relevant means from the data in Table II are: List 1, 53·2; List 2, 44·2. This difference is not significant ($p > 0{\cdot}10$), but the observation that the advantage of recognition occurs only with poorly known category members has been confirmed in a subsequent experiment (Dale, 1966).

GENERAL DISCUSSION

The results of these two experiments support the hypothesis that the beneficial effect of familiarity on free recall operates during retrieval. They concord with the theory that during retrieval subjects scan through the category of possible responses searching for those which were employed as stimuli. To summarize the evidence from the present study: formal recognition tests in which the complete category is

listed are of no advantage if the category is completely known (Experiment I), but do enhance performance with an imperfectly known category. Thus the array of alternatives on the recognition test supplements the subjects' internal list and is beneficial when this list is incomplete and where internal scanning would therefore be imperfect.

The regression analyses can be interpreted as supplying supporting evidence for this theory. Differences in familiarity between items within each list affected recall performance as indicated by the significant positive regression coefficient (cf. Table IV). They did not, however, affect recognition. Thus unfamiliar items were relatively unlikely to be recalled, but were just as likely to be correctly recognized as familiar items.

TABLE IV

Summary of Analyses Showing Mean Values with Significance Levels and the Regression of Performance on Response Availability

(Based upon split-plot analyses of variance on arc-sine transformations of the raw scores)

	Comparison	Numbers Mean	Significance	*Counties—raw scores* Mean	*Counties—raw scores* Significance	*Counties—regression coefficients* Mean values	*Counties—regression coefficients* Significance
Measure (M)	Recall	7·234	N.S.	8·006	$p < 0{\cdot}001$	+1·52	$p < 0{\cdot}01$
	Recognition	7·225		9·247		−0·03	
Test (T)	Immediate	7·650	$p < 0{\cdot}001$	8·897	$p < 0{\cdot}01$	+0·30	$p < 0{\cdot}05$*
	Delay	6·809		8·356		+1·19	
Presentation rate (R)	Rapid	6·022	$p < 0{\cdot}001$	7·806	$p < 0{\cdot}001$	+0·85	N.S.
	Slow	8·437		9·447		+0·64	
Mode of presentation (P)	Visual	7·622	$p < 0{\cdot}001$	8·516	N.S.	+0·30	$p < 0{\cdot}05$
	Auditory	6·837		8·737		+1·19	
Order (O)	1	7·022	$p < 0{\cdot}025$	8·684†	N.S.	+1·01†	N.S.
	2	7·437		8·569†		+0·48†	

* Effect entirely due to recall data.
† Here the numbers 1 and 2 refer to Lists 1 and 2, not order.

Additional evidence that when retrieving the subjects scan the complete category from which the stimuli are drawn is provided by a previous investigation. In this, two-digit numbers were used as stimuli. Analysis of the order of recall showed a tendency for subjects to write out their responses in ascending order even though no such ordering was employed during presentation (Dale and Baddeley, 1966).

In both experiments of the present study and also in the Dale and Baddeley experiment the stimuli were all drawn from a single category and the subjects were reminded during the test what that category was. It is possible that an internal search process would only occur in these special circumstances. A more general viewpoint which fits the available evidence is that recall can be regarded as problem-solving. The subject combines information from his memory traces with any other information provided by the experimenter about the constraints that exist in the experimental task. Where the stimuli are all from a single source and the subjects are reminded of this, then their problem is to match up their relevant memory traces as best they can with items known to be in the category of stimuli employed. In other situations the subjects will have to employ different clues.

In many studies in which recognition and recall performance have been compared, the recognition tests contained a very limited number of distractors so that the subject could achieve a relatively high score by guessing. When the number of distractors

is large, as in the present experiments, the value of guesswork in minimized. Nonetheless it could have had some effect upon performance. In Experiment I, a mean of 12·2 responses were made per subject in recall compared with 13·9 in recognition, while the overall mean number of correct responses in recall was 7·2. To apply a conventional guessing correction to these data it is assumed that in recognition the subjects should score 7·2 with 12 responses plus the score to be expected from two, additional, random guesses. The expected increase in this case would be less than 0·20 items. Similar calculations applied to the Counties data of Experiment II yield expected increases from guessing of 0·25 items with List 1 and 0·40 items with List 2, the greater increase arising with List 2 because of poorer recall performance.

Two points emerge from this examination of guessing correction: (1) where recognition tests yielded higher scores than recall, i.e. with List 2 of Experiment II, the advantage was much larger than the guessing correction; (2) in all other conditions the application of a correction is irrelevant since raw recognition scores were no higher than recall scores.

Despite these considerations, the possibility that recognition would have yielded higher scores than recall in Experiment I if the array had been smaller, cannot be entirely ruled out. Accordingly an additional experiment was run in which the stimuli were 12 numbers between 19 and 60. Thus the number of stimuli and the size of the array were the same as in Experiment II. Four different, randomly selected lists of stimuli were used. The presentation was auditory at a 1 sec. rate. Testing was immediate. The split group procedure of Experiments I and II was again used. A total of 66 fresh subjects, all housewives from the A.P.U. panel were employed. With recall, the mean score was 5·73 items, and the mean number of responses 9·24. With recognition, the figures were 5·70 and 10·03, respectively. Any remaining belief that the difference between the results of Experiments I and II could be attributed to the size of the array and hence to guessing must therefore be abandoned.

This research owes much to Professor B. B. Murdoch (Jr.) who encouraged and contributed to its inception, to Dr. D. E. Broadbent who first suggested that familiarity would be important, Dr. A. D. Baddeley and Dr. M. Sheldon kindly commented upon drafts. Thanks are also due to P. M. E. Altham, P. R. Freeman, and A. J. Hull, for help with statistical analyses and data processing. Dr. F. Yates kindly carried out one analysis with the computer of the Rothamsted Experimental Station.

References

Dale, H. C. A. (1966). When recognition is no better than recall. *Nature*, **211**, 324.

Dale, H. C. A. (1967). Response availability and short term memory. *J. verb Learn. verb Behav.* (in the press).

Dale, H. C. A., and Baddeley, A. D. (1966). Remembering a list of two-digit numbers. *Quart. J. exp. Psychol.*, **18**, 212–9.

Davis, R., Sutherland, N. S., and Judd, B. R. (1961). Information content in recognition and recall. *J. exp. Psychol.*, **48**, 422–9.

Poulton, E. C. (1957). Previous knowledge and memory. *Brit. J. Psychol.*, **48**, 259–70.

Tulving, E. (1964). Intratrial and intertrial retention: notes towards a theory of free recall verbal learning. *Psychol. Rev.*, **71**, 219–37.

Underwood, B. J., and Schulz, R. W. (1960). *Meaningfulness and Verbal Learning.* New York: J. B. Lippincott Co.

Manuscript received 4th January, 1967.

SHORT-TERM MEMORY AS A FUNCTION OF INPUT MODALITY

BY

SUSAN A. MARGRAIN

From Department of Psychology, Birkbeck College

The two experiments reported are concerned with short-term memory for digit lists simultaneously presented both auditorily and visually. Results showed (1) that interpolated written and verbal recall differentially affect retention depending on whether the to-be-recalled list was presented auditorily or visually. (2) That input modality appears to be far more important for recall than was directing subjects' attention to a list during input, when that list might or might not have been subsequently required for recall. The results suggest that short-term storage is modality specific. In this case, Broadbent's P and S mechanisms do not adequately describe what happens during simultaneous visual and auditory presentation. Nor would Sperling's suggestion of a final auditory store appear to be supported.

INTRODUCTION

Experiments have tended to show that auditory short-term memory is somewhat superior to visual. Broadbent (1956) found that 92 per cent. of six-digit auditory lists were correctly recalled but only 60 per cent. of similar visual lists. Buschke (1962), using simultaneous visual and auditory presentation reported that approximately 17 per cent. more repetitions of previous auditory items were recognized than previous visual items. Murray (1965) showed that at fast speeds vocalization increased recall of eight-consonant lists above that from simple reading; though more recently Murray (1966) has himself questioned this.

These results and further work using visual material (Averbach and Coriell, 1961; Averbach, 1963) led Sperling (1963) to postulate separate visual and auditory stores, although he suggests that visual material is immediately translated into subvocal input to the auditory store. Conrad (1962, 1964) supports this view. He found that errors occurring after visual presentation substantially involve acoustic confusions. However Neisser (1954) found that preliminary visual study of words facilitated the tachistoscopic recognition of them without increasing the recognition of their homonyms above new words. This suggests that effective visual storage occurred even though verbal response was required for recording all words.

Results from simultaneous presentation of visual and auditory material also suggest separate storage. Broadbent and Gregory (1961), after an earlier experiment by Moray (1960), agreed that under conditions of fast presentation digits alternately presented to either ear could pass through the P mechanism in the order of presentation but when alternated between ear and eye, they could not. They suggested that this supported Broadbent's original model involving his P and S mechanisms for simultaneous presentation; but it could be taken simply as demonstrating separate storage and not the kind of stores which might be involved.

Experiment 1 was designed to determine how far visual and auditory storage might be separate, using different orders and methods of interpolated recall after simultaneous visual and auditory presentation. Verbal recall of a visual list before recall of an auditory list, should damage recall of the latter to a greater extent than previous written recall of a visual list. This might also be expected on a theory of similar storage for both kinds of list. However, if storage was separate, written

recall of the auditory list before the visual list would damage the latter more than verbal recall of the auditory list. In this event one might ask how far Broadbent's model explained what happened during simultaneous visual and auditory presentation.

In Experiment 2 a more direct study was made of the kind of separate storage that might be involved during simultaneous visual and auditory presentation. By asking subjects to repeat either the visual or auditory items during presentation, it was possible to control which list was selected for immediate passage through Broadbent's P and which was held in S. For Broadbent's model to be supported, the difference in recall between attended-to and not-attended-to lists, regardless of input modality, should be greater than the discrepancy in recall due to input modality. If results showed the opposite, then separate modality-specific storage would be a more relevant explanation.

Experiment I

Method

Apparatus and materials. A Ferrograph tape recorder was used for simultaneous presentation of visual and auditory four-digit lists at a rate of one pair of digits per sec.; the digits being selected at random. The visual digits were stencilled on the yellow side of Basf editing tape, a ¼ in. high, and presented to pass across a 1¾ in. aperture cut in a cardboard screen, allowing a viewing time of 0·5 sec. An auditory digit was heard simultaneously. The auditory digits were recorded first, using a stop watch to regulate speed of presentation and a female voice for the recording. By stopping the tape as each auditory digit was heard, the visual ones could be added in the correct place.

Design. Twenty-four simultaneous pairs of visual and auditory lists were presented to 12 subjects (trainee teachers) in four groups of six lists each. The groups, representing the experimental conditions, differed as follows:

Group I. Written recall: auditory digits then visual digits.
Group II. Spoken recall: auditory digits then visual digits.
Group III. Spoken recall: visual digits then auditory digits.
Group IV. Written recall: visual digits then auditory digits.

N.B.—In the verbal recall condition subjects' responses were written down by the experimenter behind a screen.

Subjects, conditions and groups of lists were counterbalanced across each of the four separate parts of the test. Four practice trials were given at the start of the test using the four different experimental procedures, and subjects were rested halfway through the experiment.

TABLE I

Lists (out of Six) Correctly Recalled as a Function of Presentation Mode

	A	*V*
I. Written recall: A then V	4·90	0·67
II. Spoken „ A „ V	4·83	2·08
III. Spoken „ V „ A	2·83	4·13
IV. Written „ V „ A	3·75	4·33

Critical difference ($p < 0·05$): 1·04.
A = list presented auditorily.
V = list presented visually.

Results

A list was scored as correct if all digits were present in the correct order. Table I shows for each condition the number of lists correct out of six.

Results can be summarized as follows:—

(*a*) Delayed recall of visual lists was poor compared to delayed recall of auditory lists regardless of the recall method, even though both immediate recall conditions were similar. This supports the view that auditory retention is superior to visual in the short term.

(*b*) Although for auditory recall, conditions 3 and 4 are not significantly different, eight out of 12 subjects found more difficulty when recall was spoken. Although the difference is small, there is a tendency for interpolated spoken recall to damage auditory retention more than written interpolated recall.

(*c*) For the visual lists there is a significant difference between groups 1 and 2. This suggests that visual and auditory processing might be separate.

Furthermore it is counter to Sperling's suggestion that visual material is translated and stored in the auditory store, when one would have expected auditory interpolation to cause greater interference.

(*d*) The effect of the two interpolated recall methods, was different according to whether the recalled material was visually or auditorily presented. On Broadbent's model one might have expected interpolated auditory activity to cause more damage to both kinds of material if one assumes that it would have prevented rehearsal to a greater extent than did interpolated written activity.

One disadvantage of this design was that since the subjects knew the order of recall beforehand, they could form a set to recall in the order given. They were also free to choose which list they would attend to during presentation. In Experiment 2 only one list was required for recall, and subjects did not know which list it would be until after presentation. Furthermore by requiring subjects to speak or write one or other of the lists during presentation it was possible to manipulate which list the subject attended to, i.e. in Broadbent's terms: which list was sent immediately into P during presentation. For Broadbent's model to be supported passage through P during presentation, whether in visual or auditory form, should always result in better recall than when either list was not repeated (held in S). Therefore since verbal repetition would make the visual items effectively auditory, written repetition and spoken repetition were respectively required.

Experiment 2

Method

Materials and apparatus. The tape recorder and methods of list presentation were similar to those in Experiment 1. In addition a device was used for some conditions to permit subjects to write down a response during presentation, whilst preventing them from rereading the material they had written. Another tape recorder was used to record subjects' final recall attempts.

As in Experiment 1 simultaneous visual and auditory lists were presented, in addition these were spoken or written concurrently with presentation, to give four main conditions. Each of these main conditions was further associated with recall either of the list presented auditorily (A) or of the list presented visually (V). Each of these eight conditions included a three- four- or five-digit list presented at 1 digit/sec., and the whole test contained a replication requiring in all 48 lists.

Presentation procedure. The 48 lists were presented in four groups of 12 lists each. Whether the list was written down or verbalized depended on an ABBA sequence within

these four groups. Subjects rested for a few minutes between each group. The four main groups were subdivided into two groups of six lists each. For one group of six the visual digits, and for the other group of six the auditory digits, were all verbalized or written down during presentation. The four conditions were counterbalanced to control for fatigue and practice effects.

Recall procedure. At the end of each list the experimenter said "heard" or "seen" indicating which list was required for recall. The two types of list were recalled randomly using a different order for each subject.

Results

Lists were scored as in Experiment 1. The main results, showing for each of the eight conditions the mean number of correct lists per subject (out of two) are shown in Table II.

Table III shows the differences between conditions for combined list lengths.

TABLE II

Lists (out of Two) Correctly Recalled when Presentation Mode and List Required are Varied

Experimental conditions		*3-digit lists*	*4-digit lists*	*5-digit lists*	*Combined lists*
Presentation	*Recall*				
1. Speaks A	Reports A	2·0	1·88	1·13	1·67
2. " V	" A	1·13	1·0	0·25	0·79
3. " A	" V	1·63	0·13	0·0	0·59
4. " V	" V	1·75	1·0	0·38	1·04
5. Writes V	" V	1·50	0·50	0·13	0·71
6. " A	" V	0·50	0·38	0·0	0·29
7. " A	" A	2·0	2·0	1·13	1·71
8. " V	" A	2·0	1·63	1·38	1·67
Critical difference		0·85	0·85	0·85	0·39

TABLE III

Differences between Conditions for Combined List Lengths

	Conditions						
Condition	2	3	4	5	6	7	8
1	*0·88*	*1·08*	*0·63*	*0·96*	*1·38*	0·06	0·00
2		0·20	0·25	0·08	*0·50*	*0·92*	*0·88*
3			*0·45*	0·12	0·30	*1·12*	*1·08*
4				0·33	*0·75*	*0·67*	*0·63*
5					*0·42*	*1·00*	*0·96*
6						*1·42*	*1·38*
7							0·06

Critical difference ($p < 0{\cdot}05$): 0·39. Italics indicate significant differences.

The results show that when presentation involved auditory stimulation recall was better than when visual only, or visual-manual stimulation was involved, except for one condition in the three-digit list. Even when the visual list was written during

presentation the auditory list did better. The difference between visual and auditory recall conditions is much larger than that between lists that were written down during presentation and those that were not, although these were the lists that gained more attention during presentation. This suggests that separate storage during simultaneous visual and auditory presentation is more likely to be modality-specific than to depend on Broadbent's P and S mechanisms.

When the visual list was verbalized during presentation, recall was the same whether the recalled list was A or V. As suggested earlier this probably made the visual list effectively auditory. Furthermore presentation largely involved staggered "simultaneous" presentation of two auditory lists since subjects' verbalization came between the auditory stimuli. As Moray (1960) showed, auditory lists presented like this could be dealt with simultaneously even at this rate of presentation.

Discussion

The results of both experiments confirmed those of Broadbent (1956), Buschke (1962) and Murray (1965); i.e. auditory retention is superior to visual at least for a series of items over the short term. The similarity of these results under different presentation procedures also tends to reject the likelihood that superior auditory retention in this experiment was due to unequivalence of stimuli. Indeed Hunter and Sigler (1940) showed that visual recall was not improved by increasing the brightness, contrast and sharpness of a legible display.

The results of Experiment 1 also suggest that there may initially be two short term stores: one for visual material and one for auditory, since interpolated activity of one kind damaged the retention of earlier material of that kind, more than material of the alternative type. This occurred for both visual and auditory retention. On this modality-specific storage theory interpolated visual material could only enter and overload the visual store, and interpolated auditory material the auditory store. Any inter-modality interference could be explained by postulating as Mackworth has done (1963, 1964) that the total amount of attention is limited, and some was diverted during the presentation of interpolated material away from the critical store.

Results from Experiment 1 also indicate that separate visual storage may not be as temporary as Sperling (1963) suggested. On his model interpolated auditory activity should have damaged both visual *and* auditory retention, more than interpolated visual material. Compared to interpolated written recall, it is perhaps surprising that interpolated verbal recall was not more damaging to auditory retention than the results show.

Since results from Experiment 2 show that the difference between visual and auditory recall was much greater than that between lists repeated during presentation and lists not repeated during presentation, separate storage during simultaneous stimulation again seem more likely to be based on input modality, than on Broadbent's P and S mechanisms. As subjects were obliged to respond to both channels at different times the results cannot be explained by suggesting that the auditory list always gained access to P during presentation. Results show that when the visual list should have passed through Broadbent's P system, it was retained no more efficiently than when held in S for up to 6 sec. In fact the unrepeated auditory list did much better than the repeated visual one.

As expected when both lists involved auditory presentation (visual list verbalised) both did almost equally well. Broadbent's revised model (Broadbent and Gregory, 1961) could account for this. However, on the revised model one might have expected equally good recall from simultaneous lists that both involved visual stimulation

(auditory list written down during presentation) on the same principles as for simultaneous auditory lists.

The difference in effective storage over time for the two proposed stores reflects differences in function. Visual material is usually presented in displays of simultaneous items which are quickly resolved and replaced. Long-term retention is not required but spatial resolution is. On the other hand presentation of auditory material requires time before comprehension is complete. Consequently the auditory store must be more efficient over time.

The experiments in this paper were carried out as part fulfilment for a Ph.D. degree (thesis now at University of London Library, Senate House, W.C.1) under the supervision of Dr. J. Brown, supported by a grant from D.S.I.R. The author would also like to acknowledge the kind assistance of Dr. R. Conrad in the preparation of this paper.

References

AVERBACH, E. (1963). The span of apprehension as a function of exposure duration. *J. verb. Learn. verb. Behav.*, **2**, 60–4.

AVERBACH, E., and CORIELL, A. S. (1961). Short term memory in vision. *Bell Syst. tech. J.*, **40**, 309–28.

BROADBENT, D. E. (1956). Successive responses to simultaneous stimuli. *Quart. J. exp. Psychol.*, **8**, 145–52.

BROADBENT, D. E., and GREGORY, M. (1961). On the recall of stimuli presented alternately to the two senses. *Quart. J. exp. Psychol.*, **13**, 103–9.

BUSCHKE, H. (1962). Auditory and visual interaction in immediate memory. *J. Psychiat. Res.*, **1**, 229–37.

CONRAD, R. (1962). An association between memory errors and errors due to acoustic masking of speech. *Nature*, **193**, 1314–5.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *Brit. J. Psychol.*, **55**, 75–84.

HUNTER, W. S., and SIGLER, M. (1940). The span of visual discrimination as a function of time and intensity of stimulation. *J. exp. Psychol.*, **26**, 160–79.

MACKWORTH, J. (1963). The relation between the visual image and the post-perceptual immediate memory. *J. verb. Learn. verb. Behav.* **2**, 75–85.

MACKWORTH, J. (1964). Interference and decay in very short-term memory. *J. verb. Learn. verb. Behav.*, **3**, 300–8.

MORAY, N. (1960). Broadbent's filter theory, postulate H and the problem of switching time. *Quart. J. exp. Psychol.*, **12**, 214–20.

MURRAY, D. J. (1965). Vocalisation-at-presentation and immediate recall, with varying presentation-rates. *Quart. J. exp. Psychol.*, **17**, 47–56.

MURRAY, D. J. (1966). Effects of loudness and presentation rate on auditory short term memory. *Nature*, **210**, 226.

NEISSER, U. (1954). An experimental distinction between perceptual processes and verbal response. *J. exp. Psychol.*, **47**, 399–402.

SPERLING, G. (1963). A model for visual memory tasks. *J. Human Factors Soc.*, **5**, 19–36.

Manuscript received 12th October, 1966.

EFFECTS OF VARYING THE APPARENT TILT OF THE INSPECTION FIGURE ON A KINAESTHETIC AFTER-EFFECT

BY

R. OVER

From the Psychology Department, University of Otago

Three experiments were conducted to determine whether kinaesthetic after-effects, measured by settings of a bar to the horizontal, are controlled by the apparent or the physical tilt of the bar during inspection. Two arrangements were used to induce a discrepancy between apparent and physical tilt. Data obtained with one arrangement did not permit a proper test of the issue. With the other arrangement either distorted or undistorted visual information about tilt was present during kinaesthetic inspection. It was found that post-inspection settings were displaced from pre-inspection settings in the direction of the tilt of the inspection figure when there was no discrepancy between information presented by the two modalities. When there was a discrepancy, the displacements attributable to apparent tilt were in the direction opposite to the apparent tilt of the inspection figure. The former effects dissipated rapidly after inspection; the latter did not. It is suggested that after-effects obtained following an inspection period during which there is no discrepancy between physical and apparent tilt are controlled by the physical tilt of the inspection figure and that apparent tilt, when it is an effective variable, operates by modifying the judgemental frame of reference.

INTRODUCTION

The present experiments consider whether kinaesthetic after-effects, measured by requiring the subject to make judgements of tilt, are influenced by the physical or by the apparent spatial properties of the inspection figure. A number of studies (cf. Sutherland, 1961) have asked whether visual after-effects involving size judgements are controlled by the apparent size of the inspection figure. Two types of stimulus arrangements have been employed. Either the inspection figure (IF) and the test figure (TF) have been of the same physical size but at different physical distances from the subject, or IF and TF have been unequal in physical size but have subtended equal visual angles. It has been argued that in the former case IF and TF are equal in apparent size but unequal in retinal size, while in the latter case they are unequal in apparent size but equal in retinal size. The relative effects of the two classes of variables have been determined by contrasting the judgements found with the two arrangements. The results from experiments which have employed this design are not unequivocal; a summary and discussion of many of the studies is provided by Sutherland (1961).

Gibson (1937), Gibson and Backlund (1963), and Day and Singer (1964) have reported a kinaesthetic after-effect which involves judgements of orientation. Following kinaesthetic inspection of a bar tilted 15° in a clockwise direction from the horizontal, a bar which is physically horizontal feels to be tilted about 5° in an anti-clockwise direction. Judgements of orientation made before and after inspection do not differ when the bar is physically horizontal during inspection. To determine whether kinaesthetic after-effects are controlled by the apparent or the physical tilt of the bar during inspection it is necessary to arrange experimental conditions such that the bar, when at a given physical tilt during inspection, either feels to the subject to be at that tilt (same physical and apparent tilt) or feels to be at some other tilt (different physical and apparent tilts).

Experiment I

An intersensory conflict situation is used in Experiment I to provide the above conditions. When visual and kinaesthetic information differ it has been found (Gibson, 1933; Rock and Victor, 1964; Over, 1966) that judgements of the felt spatial properties of an object are controlled by its visible spatial properties. Thus if a subject is required to move his hand across a horizontal bar while viewing through prisms which make the bar look to be tilted 15° from the horizontal, he reports that the bar feels to be tilted in accord with its visual tilt (Over, 1966). By these means it is possible to present, during inspection, a bar which is physically horizontal but which looks and feels to the subject to be tilted 15° from the horizontal. If after-effects are controlled by the physical slant of IF, settings made in the absence of vision after inspection should not differ from those made before inspection. If, however, after-effects are controlled by the apparent slant of IF, a difference should be found between post- and pre-inspection settings. The displacement should be in the direction of the apparent slant of IF. It is also possible to arrange conditions such that a bar which is physically tilted 15° from the horizontal looks and feels to be horizontal during inspection. Table I sets out the differences between post- and pre-inspection settings expected by the "apparent tilt" and the "physical tilt" hypotheses for this and the earlier condition, as well as for the conditions where visual information is either not available during inspection or does not conflict with kinaesthetic information. These latter arrangements constitute the control conditions of Experiment I; it is assumed that there is no difference between the apparent tilt and the physical tilt of IF under these conditions.

TABLE I

Predictions from the "Apparent Tilt" and the "Physical Tilt" Hypotheses

Inspection condition	*"Apparent tilt" hypothesis*	*"Physical tilt" hypothesis*
Ko	No	No
K15	Yes	Yes
Ko Vo	No	No
Ko V15	Yes	No
K15 Vo	No	Yes
K15 V15	Yes	Yes

Ko indicates that IF was physically horizontal; K15 that IF was tilted 15° from the horizontal; Vo that IF looked horizontal; and V15 that IF looked to be tilted 15° from the horizontal.

"No" indicates that the particular hypothesis predicts that no after-effect (that is, no difference between post- and pre-inspection settings) will be obtained under the given inspection conditions; "yes" that an after-effect will be found.

Method

Subjects. There were six groups, each of 10 subjects. All subjects were students beginning an introductory course in psychology.

Apparatus. A pivoted wooden bar, 24 in. long, 3·5 in. wide, and 1 in. thick, was set 14 in. in front of an eyepiece and could be seen by the subject only when viewed through the eyepiece. Two Dove prisms, 0·6 in. × 0·9 in. × 1·88 in., were mounted in the eyepiece; one of these could be rotated by the experimenter so that the bar could be made to look to be at any given tilt. When the bar both was and looked to be horizontal the bottom half of the subject's field of view was filled by the bar, which was black, and the top half by a white background. The subject was able to place his right hand on the bar

by reaching beneath the framework which supported the eyepiece. He was required to move his hand along the centre 6·25 in. section of the bar, the limits of which were marked by transparent tape; this section corresponded to the limits of the horizontal field. Movements were made in time with a metronome operating at 84 beats/min.; because the subject's arm was not supported, movements of the hand, arm, and shoulder were involved. The subject was able to change the tilt of the bar, when required, by operating rotary controls with his left hand. Settings of the bar were read by the experimenter to the nearest 0·25° from a protractor and marker mounted at the rear of the frame supporting the bar.

Procedure. Nine pre-inspection settings were initially obtained from each subject. The subject was required to move his hand across the bar in time with the metronome and to complete, between 10 and 15 sec., a setting of the bar so that it felt horizontal. A shutter prevented the subject from seeing the bar while making these settings. Three of the settings were obtained from each of three starting positions—15AC (where the bar was tilted 15° in the anti-clockwise direction from the horizontal), 15C (15° in the clockwise direction), and H (horizontal). The mean of these nine settings was used as the pre-inspection setting with which the post-inspection settings obtained later were compared to determine amount of after-effect.

Each subject made nine post-inspection settings; these were also made in the absence of visual information about the tilt of the bar. Each post-inspection setting was immediately preceded by an inspection period of 90 sec. during which the subject was required to move his hand across the bar in time with the metronome. The six groups differed in terms of the information, relating to the tilt of the bar, with which they were provided during inspection. The arrangements used for the different groups are summarized in Table I, together with the predictions from the two hypotheses. The tilt of the bar was in the clockwise direction for half of the subjects, and in the anti-clockwise direction for the other half, under each condition where the bar either was or looked to be tilted during inspection. Subjects in groups which were provided with visual information were carefully watched to ensure that they kept their eyes open during the entire inspection period. For these subjects a shutter was placed in front of the eyepiece at the end of inspection. Subjects in groups K0 and K15 never saw the bar at any stage of testing. Immediately after each inspection period each subject was required to make one post-inspection measure by setting the bar so that it felt horizontal. Nine post-inspection measures were obtained in all from each subject; there were three from each of the three starting positions (15AC, 15C, H). A rest of 4 min. was given between each inspection period.

Results and Discussion

Table II shows the mean differences found between post- and pre-inspection settings of the six groups, as well as the standard deviations. The score + has been used to signify that post-inspection settings were displaced from pre-inspection settings in the direction of the physical tilt of IF, if the bar was tilted during inspection,

TABLE II

MEANS AND S.D.S OF DIFFERENCES BETWEEN POST- AND PRE-INSPECTION SETTINGS

Inspection condition	*Mean*	*S.D.*	*t*
K0	0·06AC	1·47	0·00
K15	+3·76	1·51	7·47*
K0 V0	0·18AC	1·22	0·33
K0 V15	−2·02	0·88	6·89*
K15 V0	+3·19	0·78	13·87*
K15 V15	+4·85	0·96	15·20*

* $p < 0·01$.

or the visual tilt of IF, if the bar was horizontal during inspection. For groups Ko and Ko Vo, for which after-effects are not predicted by either hypothesis, mean differences have been scored as AC (anti-clockwise) or C (clockwise) displacements.

As expected, the mean differences obtained under conditions Ko and Ko Vo do not differ significantly from zero. Displacements differing from zero in the expected direction have been found under conditions K15 and K15 V15; for subjects in these groups the bar both was and felt to be tilted during inspection. Under conditions Ko V15 and K15 Vo there was a discrepancy between the physical and the apparent tilt of the bar during inspection. Displacements differing significantly from zero were obtained under both conditions. When the bar was horizontal during inspection but felt tilted, post-inspection settings were displaced in the direction opposite of the tilt of IF. When IF was tilted but felt horizontal, post-inspection settings were displaced both in the direction of the physical tilt of IF and in the direction opposite to the visual distortion. The mean displacement under condition K15 V15 is significantly greater ($t = 4{\cdot}02$, $p < 0{\cdot}01$) than that found under condition K15 Vo.

It can be seen that post-inspection settings are influenced by both the physical and the apparent tilt of IF when these two variables are different during inspection. The displacements produced by the apparent tilt differ, however, from those produced by the physical tilt. The bar was reported to feel tilted slightly in a clockwise direction from the horizontal following inspection during which IF both was and felt to be tilted in a clockwise direction. When IF was horizontal but felt to be tilted in a clockwise direction, the subsequent displacement was in an anti-clockwise direction. The former finding is consistent with a treatment of after-effects in terms of a process of adaptation to a spatial norm. Gibson (1937, 1959) has claimed that, for each sensory quality which falls into an opposition series, adaptation occurs when there is inspection of a stimulus property which evokes a quality removed from the neutral point. During inspection this quality becomes more neutral—the bar comes to feel less tilted—and immediately after inspection the covariation of the various stimuli of the dimension and the various qualities of the continuum will remain correspondingly modified. The latter finding is consistent with the claim (Harris, 1965) that an after-effect obtained following a period of adaptation to displaced vision results from the arm coming, during inspection, to feel to be where it looks to be. Thus a bar which is horizontal but looks to be tilted in a clockwise direction will feel to be tilted clockwise during inspection. When judgements are subsequently made in the absence of vision a physically horizontal bar will feel to be tilted in a clockwise direction.

Experiment II

It has been suggested that both normalization after-effects, controlled by the physical slant of IF, and prism-induced after-effects, found when there is a discrepancy between the physical and the apparent tilt of IF, have been obtained in Experiment I. There has been some speculation as to the relationship between these two classes of effects. Gibson (1959) has suggested that both arise from adaptation to spatial norms. It has been shown in Experiment I, however, that the two effects differ in terms of their direction. Further differences between the two classes of displacements are suggested from data from other sources. After-effects obtained when there is no discrepancy between the physical and the apparent tilt of IF are known to dissipate rapidly after inspection (Singer and Day, 1965), while after-effects found when there is a difference between the physical and the apparent location of IF dissipate slowly (Hamilton and Bossom, 1964).

Experiment II compares the rates of dissipation of after-effects obtained following inspection under conditions K15, Ko V15, and K15 Vo. By determining whether after-effects obtained under conditions Ko V15 and K15 Vo disspiate at the same rate as those found under conditions K15, it can be asked whether one or two sets of processes underlie the effects obtained in Experiment I.

Method

Subjects. There were three groups, each of 12 subjects. The subjects were undergraduate students.

Apparatus. The equipment was the same as that used in Experiment I.

Procedure. The treatments of the three groups differed only in terms of the information available during inspection. The three conditions used—K15, Ko V15, and K15 Vo—are as described in Experiment I. Nine pre-inspection settings were initially obtained from each subject. Each subject made 12 post-inspection measures, each of which was preceded by an inspection period of 90 sec. duration. Three post-inspection settings were made at each of four delay times after the end of inspection—0, 30, 60 and 120 sec. When the delay was 0 sec. the subject made a setting immediately following the end of the inspection period. In testing at the other delay periods the subject removed his hand from the bar at the end of the inspection period and placed it flat on his knee until asked to make a setting. The subject was requested to make each setting within 20 sec. of placing his hand back on the bar. For each delay period one setting was obtained from each of the three starting positions (15AC, 15C, H). Possible cumulative effects were counterbalanced by varying the order of presentation of the delay periods by a Latin square arrangement. A rest period of 6 min. was given between the end of one inspection period and the commencement of the next.

Results and Discussion

The mean differences between post- and pre-inspection setting for the three groups and the four delay periods are shown in Table III, together with the standard deviations. The scoring notation used is that employed in Experiment I. The analysis of variance summarized in Table IV shows that the dissipation of the displacement produced by inspection is a function of the inspection conditions. The after-effect obtained under K15 conditions has almost entirely dissipated 60 sec. after inspection; the mean displacement obtained at that delay is not significantly different from zero. The effects which follow inspection during which there is a discrepancy between the apparent and the physical tilt of IF do not, however, dissipate rapidly; for each of the groups tested under these conditions the mean displacement is significantly different from zero at each delay time.

TABLE III

MEAN DISPLACEMENTS AT DIFFERENT TIMES AFTER INSPECTION

Inspection condition		*Time after inspection (sec.)*			
		0	*30*	*60*	*120*
K15	Mean	+4·43	+2·44	−0·09	+0·98
	S.D.	1·44	1·32	3·49	1·95
Ko V15	Mean	−1·54	−2·22	−2·50	−2·00
	S.D.	1·67	2·19	2·33	2·01
K15 V15	Mean	+5·11	+4·15	+4·22	+3·28
	S.D.	1·59	2·23	3·18	1·26

TABLE IV

TEST OF TREND OF DISPLACEMENTS AT DIFFERENT TIMES AFTER INSPECTION

Source	*d.f.*	*MS*	*F*
Inspection conditions (I) ..	2	398·96	27·96*
error (*a*)	27	14·27	
Delay (D)	3	24·23	11·01*
I × D	6	7·61	3·46*
error (*b*)	81	2·20	
Total	119		

* $p < 0.01$.

It is known from Experiment I that post- and pre-inspection settings do not differ when IF both is and appears to be horizontal. Displacements are found both when IF is and feels to be tilted 15° and when IF is horizontal but feels to be tilted 15°. The present experiment has confirmed that these displacements differ in direction; it has also shown that they differ in rate of dissipation. These data indicate that after-effects obtained under conditions where there is no discrepancy between the apparent and the physical tilt of IF are controlled by the physical tilt of IF.

The data also show that the normalization after-effects reported by Gibson (1933) and prism-induced after-effects (Harris, 1965) are controlled by different sorts of processes. The former dissipate rapidly like many sensory effects which are dependent on the duration of inspection; they thus can be called sensory after-effects and distinguished from prism-induced displacements, which can be termed perceptual after-effects. It is probable that the displacements obtained under conditions Ko V15 and K15 Vo represent the combined effects of these two classes of processes. Inspection conditions K15 Vo, for example, may induce a sensory after-effect of about 4° and a perceptual after-effect of 2°. The former would be in the direction of the physical tilt of IF; the latter in the direction opposite to the visual distortion. When the bar is tilted but looks to be horizontal, these two effects would be cumulative and the decreases in displacement over delay times would represent the dissipation of the sensory after-effect.

An issue of some importance to theories of sensory after-effects can be examined by a closer study of the interaction of sensory and perceptual after-effects. Experiment I showed that post- and pre-inspection settings did not differ under condition Ko. It can be asked whether this indicates that inspection of a horizontal bar produces no after-effect or an after-effect of 0°; that is, whether inspection of a horizontal bar produces sensory effects which are qualitatively similar to, but quantitatively different from, sensory effects produced by inspection of a tilted bar. If a sensory after-effect of 0°, rather than no sensory after-effect, is produced by inspection of a horizontal bar, displacements obtained immediately after inspection of a horizontal bar which looked tilted would represent the interaction of a sensory after-effect of 0° and a perceptual after-effect of about 2°. They thus should be less than 2°. Some time after inspection, however, the sensory after-effect would have dissipated and the displacement would be determined by the perceptual after-effect alone. The data obtained under conditions Ko V15 are not sufficiently precise to examine this possibility; although the after-effect obtained immediately after inspection is less than the after-effects found at the other delay times the differences

are not significant ($F = 1{\cdot}49$, $p > 0{\cdot}05$). It is worth noting, however, that the problem, which relates to the sorts of processes which underlie sensory after-effects, can be studied only within an experimental situation in which the conditions necessary for both classes of after-effect are present.

Experiment III

Experiment III employs a different arrangement to produce a discrepancy between the apparent and the physical tilt of IF. Some subjects, when required in pre-inspection settings to set the bar to the horizontal, consistently set the bar at an anti-clockwise tilt relative to the physical horizontal, and others at a clockwise tilt. It can be assumed that a physically horizontal bar presented as IF will feel to be tilted by these subjects. If after-effects are controlled by the physical tilt of IF, post-inspection settings made by these subjects should not differ significantly from pre-inspection settings. If, however, after-effects are controlled by the apparent tilt of IF, post-inspection settings should be displaced from pre-inspection settings in the direction of the apparent tilt of IF.

Method

Subjects. The subjects were undergraduate students. A number of students were tested to form the two experimental and the two control groups, each of which contained eight subjects. The criterion used to select subjects is outlined later.

Apparatus. The equipment was the same as that used in Experiments I and II. At no time during testing were the subjects able to see the bar.

Procedure. The initial aim was to select subjects for whom there was a consistent discrepancy between the apparent horizontal and the physical horizontal. People in the initial pool of volunteers were each required to make eight settings of the bar so that it felt horizontal. Two starting positions of the bar—15AC and 15C—were used, and four settings were made from each position. For each experimental and control group half of the subjects made settings in the order 15AC–15C, and half in the order 15C–15AC. Those people who misaligned the bar in a consistent direction on all eight settings were selected as subjects for further testing. The other people were not tested any further.

Subjects who set the bar at an anti-clockwise tilt in their pre-inspection settings were allocated to either an experimental or a control group, each of which contained eight subjects, as were subjects who set the bar at a clockwise tilt. Each experimental subject made two post-inspection settings, each of which was preceded by an inspection period of 90 sec. During inspection the subject was required to move his hand along the bar, set at the physical horizontal, in time with a metronome operating at 84 beats per min. For each subject the first post-inspection setting was made from the starting position used to obtain the first four pre-inspection settings, and the second from the other starting position. There was a rest period of 5 min. between the two inspection periods.

Subjects in the control groups each made two settings, neither of which was preceded by an inspection period. The second setting was made 6·5 min. after the first. Control groups were used to determine whether subsequent settings differ from pre-inspection settings as a function of the time between making settings rather than the spatial properties of IF.

Results and Discussion

The means and standard deviations of the pre-inspection measures are given in Table V for the four groups. These values show the extent to which subjects initially misjudged the horizontal. They indicate that the physically horizontal bar present during inspection would have felt to be tilted in a clockwise direction by subjects in group E/AC and in an anti-clockwise direction by subjects in group E/C. The mean differences found for these groups by subtracting pre-inspection values from the post-inspection setting made from the same starting position are shown in Table V,

together with the standard deviations. The score — signifies that post inspection settings are displaced from pre-inspection settings in the direction of the apparent tilt of IF. It should be noted that the score — also indicates that post-inspection settings were more accurate, relative to the physical horizontal, than were pre-inspection settings. Mean differences between subsequent and prior settings made by the control groups are also shown in Table V. For these groups the score + has been used to indicate that subsequent settings were more accurate, relative to the physical horizontal, than prior settings.

TABLE V

MEAN PRIOR SETTINGS AND DIFFERENCES BETWEEN SUBSEQUENT AND PRIOR SETTINGS

Group		*Prior settings*		Differences
		15AC	*15C*	
C/AC	Mean	3·05AC	3·16AC	+1·68
	S.D.	1·17	1·12	1·27
C/C	Mean	3·48C	3·28C	+1·51
	S.D.	1·50	1·58	1·12
E/AC	Mean	3·60AC	3·41AC	+2·16
	S.D.	2·56	1·30	0·58
E/C	Mean	3·55C	5·05C	+2·32
	S.D.	2·61	2·12	2·29

C/AC signifies control subjects who initially set the bar in an anti-clockwise direction from the horizontal, and C/C control subjects who initially set the bar in a clockwise direction from the horizontal. E refers to experimental subjects.

For each group the mean differences are in the + direction and are significantly different from zero. The mean displacements obtained for groups C/AC and E/AC do not differ significantly ($t = 0{\cdot}91$, $p > 0{\cdot}05$); nor do those found for groups C/C and E/C ($t = 0{\cdot}84$, $p > 0{\cdot}05$). The results of the control groups indicate that subjects who are initially inaccurate in their judgements of the horizontal become more accurate over trials and over time. It is probable that the results of the experimental groups represent these changes rather than the effects of the inspection conditions. Subjects for whom the discrepancy between the apparent horizontal and the physical horizontal is stable over a long period of time need to be used before the issue can be properly examined. Even with such subjects constant errors in judging the horizontal may have to be large before after-effects would be expected following inspection of a physically horizontal bar.

General Discussion

Two experimental arrangements have been used to produce a discrepancy between the apparent and the physical tilt of IF, and thereby to determine which variable influences subsequent settings of TF. The results obtained under the control conditions in Experiment III suggest that the post-inspection settings of the experimental subjects reflect a trend towards greater accuracy over trials rather than the effects of the inspection conditions. This arrangement thus has not permitted a proper test of the issue. In Experiments I and II relationships between the apparent and the physical tilt of IF were systematically varied by controlling the visual

information available to the subject during inspection. Available data (Gibson, 1933; Over, 1966; Rock and Victor, 1964) indicate that visual information influences where IF feels to be; that is, its apparent tilt. After-effects obtained when visual and kinaesthetic information was congruent during inspection did not differ from those found when kinaesthetic information alone was available during inspection. When visual information was distorted, after-effects were found to be a function of the apparent, as well as the physical, tilt of IF. The after-effects controlled by the two classes of variables differed, however, in their direction and rates of dissipation. These differences are such that after-effects obtained when there is no discrepancy between the apparent and the physical tilt of IF cannot be attributed to the apparent tilt of IF.

In several studies (Day and Logan, 1961; Story, 1961; Sutherland, 1954) it has been found that the direction of a visual after-effect involving size judgements is determined by the relative apparent sizes of IF and TF when these figures subtend equal visual angles. The present experiments suggest an approach by which it can be determined whether the above after-effects are equivalent to those obtained under conditions where the apparent and the retinal image size relationships of IF and TF are confounded as variables. The rate of dissipation of after-effects obtained under the latter conditions is a function of the duration of the inspection period (Oyama, 1953). If after-effects determined by apparent size relationships alone dissipate at the same rate it can be concluded that after-effects found when apparent and physical size relationships are confounded as variables are in fact controlled by apparent size relationships. If they do not dissipate, or if they dissipate at a different rate, it follows that the after-effects found when the two variables are confounded are determined by retinal image size relationships. Results of the latter sort would indicate that it is profitable to distinguish between sensory and judgemental determinants of judgements of TF following inspection.

The issue of intermodal after-effects can be examined along the same lines. Jaffe (1956) and Mayer (1961) have reported data which suggest that inspection in one modality (vision) results in after-effects in another modality (kinaesthesis). If these intermodal after-effects are equivalent to after-effects found when inspection and subsequent test settings are accomplished within the one modality, both should dissipate at the same rate following inspection. If they do not, it can be said that prior stimulation in one modality affects judgements made in another, but that the processes involved are quite different from those operating when judgements made in a modality are influenced by prior inspection involving that same modality.

References

Day, R. H., and Logan, J. A. (1961). A further investigation of apparent size and retinal size as determinants of the figural after-effect. *Quart. J. exp. Psychol.*, **13**, 193–203.

Day, R. H., and Singer, G. (1964). Spatial after-effects within and between kinesthesis and vision. *J. exp. Psychol.*, **68**, 337–43.

Gibson, J. J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *J. exp. Psychol.*, **16**, 1–31.

Gibson, J. J. (1937). Adaptation with negative after-effect. *Psychol. Rev.*, **44**, 222–44.

Gibson, J. J. (1959). Perception as a function of stimulation. In Koch, S. (Ed.), *Psychology: A Study of Science*. Vol. 1. New York: McGraw Hill.

Gibson, J. J., and Backlund, F. A. (1963). An after-effect in haptic space perception. *Quart. J. exp. Psychol.*, **15**, 145–54.

Hamilton, C. R., and Bossom, J. (1964). Decay of prism aftereffects. *J. exp. Psychol.*, **67**, 148–50.

HARRIS, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychol. Rev.*, **72**, 419–44.

JAFFE, R. (1956). The influence of visual stimulation on kinesthetic after-effect. *Amer. J. Psychol.*, **69**, 70–5.

OVER, R. (1966). An experimentally induced conflict between vision and proprioception. *Brit. J. Psychol.*, **57**, 335–342.

OYAMA, T. (1953). Experimental studies of figural aftereffects: I. Temporal factors. *Jap. J. Psychol.*, **23**, 239–54.

ROCK, I., and VICTOR, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, **143**, 594–6.

SINGER, G., and DAY, R. H. (1965). Temporal determinants of a kinesthetic aftereffect. *J. exp. Psychol.*, **69**, 343–8.

STORY, A. W. (1961). Has apparent size been tested as a factor in figural after-effects? *Quart. J. exp. Psychol.*, **13**, 204–8.

SUTHERLAND, N. S. (1954). Figural after-effects, retinal size and apparent size. *Quart. J. exp. Psychol.*, **6**, 35–44.

SUTHERLAND, N. S. (1961). Figural after-effects and apparent size. *Quart. J. exp. Psychol.*, **13**, 222–8.

Manuscript received 12th July, 1966.

REACTION TIME TO A SECOND STIMULUS AS A FUNCTION OF INTENSITY OF THE FIRST STIMULUS

BY

MARILYN C. SMITH

From Department of Psychology, University of Toronto

Reaction time (RT) to the second of two stimuli presented in rapid succession was examined as a function of the intensity of the first stimulus (S_1). It was found that the delay in RT_2 was greater following a dim first stimulus than following a bright first stimulus. The magnitude of this increase corresponded to the difference in RTs to the two intensity levels of S_1. These results support the prediction of a single channel model of response selection. Examination of mean first RTs revealed a general elevation in latency of RT. However, since this increase was not influenced by the inter-stimulus interval (ISI) or by the intensity of the second stimulus (S_2), and since the same increase was found on "catch trials" where no S_2 was presented, this increase is considered to be a function of change in set in the double response situation.

Introduction

When two stimuli are presented in rapid succession, the reaction time (RT) to the second stimulus is typically delayed when compared with the RT to that stimulus when it is presented alone. This delay, which is commonly referred to as the "Psychological Refractory Period," is maximal when the inter-stimulus interval (ISI) is very short, and declines as the interval between the two stimuli is increased.

One explanation of this delayed second RT is that somewhere in the processing system there is a "single channel" or limited capacity mechanism which can handle the information from only one of the stimuli at a time. This single channel is usually considered to occur at the decision or response selection stage (Hick, 1948; Craik, 1948; Welford, 1952; Davis, 1956; Fraisse, 1957). A single channel model of the psychological refractory period attributes the delay in the second RT to the fact that the information from the second stimulus must be held in store until the single channel decision mechanism has completed the first response selection.

According to this model, the delay in the second RT (RT_2) is a function of the time required to select the first response. It would therefore follow that the sooner the first response selection is begun, the sooner it will be completed (provided the selection time is constant), and therefore the sooner the single channel will be freed for the second response selection. Consequently, the delay in RT_2 should be reduced.

Our experiment was conducted to examine this prediction. The time of initiation of the response selection was manipulated by varying the intensity of the first stimulus (S_1). It has long been known that RT to a stimulus decreases as the intensity of the stimulus is increased. Recently, Vaughan, Gilden and Costa (1964) have demonstrated that the latency of the visual evoked response varies inversely with the intensity of the stimulus. Hence, increasing the intensity appears to operate primarily by reducing the time required for the stimulus information to reach the cortex, and therefore the time at which response selection can begin.

It was hypothesized that if increasing the intensity of a stimulus reduced the RT to that stimulus by x millisec., this was due to the fact that response selection was begun, and therefore completed, x millisec. sooner. Consequently if this bright stimulus were to be used as the first stimulus in a situation where two stimuli were

presented in rapid succession, the response selection to S_2 should be able to start x millisec. sooner than if it follows a dim S_1, with a resulting decline in RT_2 by x millisec.

A second purpose of this experiment was to examine changes in the RT to the first stimulus (RT_1). Previous findings on the changes in RT_1 in a successive response situation have been contradictory. Reynolds (1966) has reported a general rise in mean first RT as a function of ISI length, which he explains on the basis of psychophysical summation at short ISIs. Others, however, have found no change in RT_1 (Davis, 1956), or a slight increase in RT_1 which is the same across all ISIs (Smith, 1966). Gottsdanker, Broadbent and Van Sant (1963) compared mean first RT at an ISI of 500 millisec. with the choice RT to that stimulus presented alone, and found the former to be significantly larger. In this study, changes in RT_1 could be examined as a function both of ISI and of intensity of S_2.

Method

Subjects

Four M.I.T. students, three female and one male, participated in this experiment. The subjects were paid for their services.

Apparatus

(*a*) *Visual array.* The visual array was composed of two light sources consisting of cold-cathode mercury-argon lamps coated with magnesium tungstate phosphor. The lamps were housed in circular light boxes having a diffusing surface made of $\frac{1}{4}$ in. milk plexiglass. A small red focussing light was placed between these two stimuli. The area of each of the circular light sources was 76 $cm.^2$ The visual angle of each of the light sources was 3·8 degrees, while that of the entire array was 12·4 degrees. These lights were triggered by a Sky Instrument Company type tachistoscopic programmer which enables the interval between the stimuli to be varied. The lights could be triggered in either order—i.e. left followed by right, or right followed by left.

(*b*) *Reaction time apparatus.* The RT keys consisted of two microswitches mounted on a handrest. A switch operated by the experimenter sounded a buzzer and activated a model 111C Hunter decade interval timer. The timer, after an interval of 1·4 sec., triggered the simultaneous onset both of the first light and of two model 120A Hunter Klock-kounters which measured RT in millisec. Pressing the key to the first light stopped the first timer; pressing the key to the second light stopped the second timer. A chin-rest was provided for the subject's head, to prevent head movements and to facilitate constant direction of gaze.

Procedure

The experiment was conducted in a dimly-lit room. The subject faced the visual array and focussed on the red light between the two stimulus lights. His left fore-finger was placed on the left microswitch and his right forefinger rested on the right microswitch. The experimenter began each trial by saying "Ready" and pressed a buzzer. Following the buzzer there was a fixed foreperiod of 1·4 sec., which was followed by the onset of one of the two lights. On half the trials the left stimulus came on first, and on the other half the right stimulus came on first. The left-right order was randomly varied. The interval from the onset of the first stimulus to the onset of the second stimulus (ISI) was varied in 30 millisec. steps in a random fashion from 60 to 270 millisec. The duration of both stimuli was held constant at 30 millisec. A summary of the procedure is as follows:

"Ready" ------ Buzzer --- 1·4 sec. --- Stimulus one ------ Variable interval ------ Stimulus two

30 millisec. 30, 60 --- 240 millisec. 30 millisec.

ISI = 60, 90, --- 270 millisec.

The subject was instructed to respond as quickly as possible, pressing the left key when the left light came on and the right key when the right light came on. Each session

lasted approximately 1 hr. Subjects attended 15 sessions, of which only the last 10 were included in the data analysis.

Design

Two luminance levels were used, 42 ft. L. (Bright) and 0·042 ft. L. (Dim). The luminance was reduced by lowering a Kodak No. 3 neutral density filter into a slot in front of the light source by means of a pulley arrangement. The filters were adjusted on each trial, so that the subject did not have any cues as to whether the intensities would be the same as or different from the previous trial. Four luminance level combinations were possible:

Stimulus one	*Stimulus two*
Bright	Bright
Bright	Dim
Dim	Bright
Dim	Dim

Each of the above luminance levels could appear either with the first stimulus on the left or the first stimulus on the right, giving eight possible luminance combinations. For each of these eight combinations, all eight ISIs were presented, making a total of 64 trials. In addition, 16 "catch-trials" were included—i.e. trials in which only the first stimulus was presented, to insure that the subject would not merely press both keys in sequence as soon as he saw the first stimulus. This also provided a measure of his RT to one stimulus when he believed it to be the first of two, to see if this differed from the choice reaction to one stimulus when he knew only one was being presented. Half the catch-trials were with the bright stimulus and half with the dim stimulus. The side of presentation was counterbalanced.

Finally, 36 simple reaction time (SRT) and 36 choice reaction time (CRT) measures were taken—one-third at the beginning of the session, one-third in the middle of the session and one-third at the end. A constant foreperiod of 1·4 sec. was used. These measures provided a measure of how RT to a single stimulus was affected by luminance level. In addition, the choice RT measure allowed us to determine if RT_1 in the double stimulation situation was different from the two-choice RT to a single stimulus.

Results

A. *Effect of intensity level on RT*

The effect of the luminance of the visual stimulus on RT when only a single stimulus was presented is shown in Table I. These results confirm earlier findings

TABLE I

RT in Millisec. to a Visual Stimulus at Two Intensity Levels

Subjects	*Simple RT*			*Choice RT*		
	Intensity of stimulus		*Difference*	*Intensity of stimulus*		*Difference*
	Dim	*Bright*		*Dim*	*Bright*	
1	199	175	24	228	196	32
2	226	201	25	277	247	30
3	223	192	31	255	230	25
4	232	204	28	261	228	33
Average	220	193	27	255	225	30

Each entry represents the mean of 180 trials.

that when a subject is required to respond to stimuli of different luminance levels, RT is faster to the stimulus of greater intensity. With the particular intensities used in this experiment, RT is approximately 30 millisec. faster to the bright light than to the dim light. This is true for both simple and choice RT. To determine whether this difference is found as well in the double stimulation situation, the mean first RTs under all conditions were examined. These are presented in Table II.

TABLE II

MEAN FIRST REACTION TIME IN MILLISEC.

		ISI in millisec.									
		60	*90*	*120*	*150*	*180*	*210*	*240*	*270*		*Average*
S_1 bright	S_2 bright	274	267	274	266	267	261	269	274		269
	S_2 dim	258	270	259	263	264	267	271	264		264
										Mean	267
S_1 dim	S_2 bright	287	301	295	303	304	308	305	304		301
	S_2 dim	296	297	297	296	299	307	304	295		299
										Mean	300

Each entry represents the mean of 20 trials.

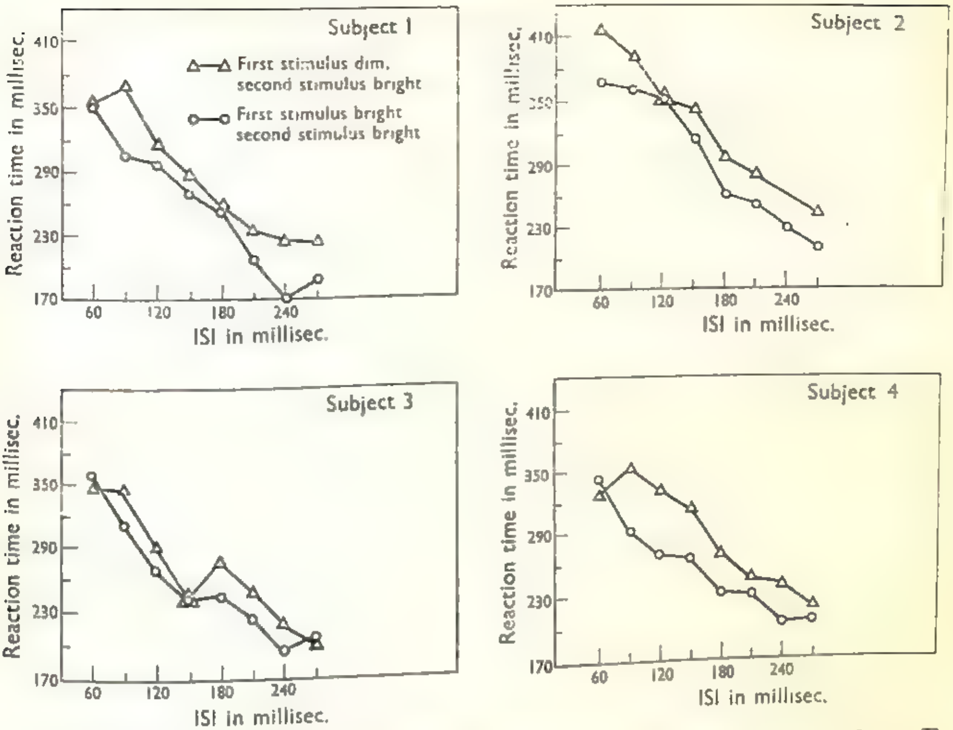
Examination of the last column in Table II shows that the mean first RT to a bright stimulus (267 millisec.) is approximately 30 millisec. less than the mean first RT to a dim stimulus (300 millisec.).

B. Effect of the intensity of the first stimulus upon RT to the second stimulus

When the first stimulus was bright and the second stimulus dim, subjects often reported being unable to see the second stimulus, and assumed that this was a catch-trial on which only one stimulus was presented. Responses which were made to the second stimulus were abnormally long. Possibly the intense first stimulus has a masking effect on the second dim stimulus. Consequently, examination of the effect of the intensity of a first stimulus on the RT to a second stimulus of a given intensity was limited to those trials on which the second stimulus was bright. The results for the four subjects are shown in Figure 1. Each point represents the mean of 20 trials. For all four subjects RT to the second stimulus is less if the preceding stimulus was bright than if it was dim. The only exception to this occurs at the shortest ISI employed, in the case where the first stimulus was dim and the second stimulus was bright. Under these conditions the effective ISI between the two stimuli was considerably less than 60 millisec. (when the relative rates of travelling of the two intensities of stimuli are considered), and some "grouping" may have occurred, such that the two stimuli were processed simultaneously, and a single response selection was made to the stimulus array.

The mean difference in RT to the bright stimulus when it is preceded by a dim S_1 and when it is preceded by a bright S_1 is shown in Table III. This is compared with the mean difference in first RTs to the two intensity levels of the first stimulus. In both cases the value of this difference is 30 millisec. It thus appears that if the second stimulus is preceded by a dim first stimulus, RT is increased by 30 millisec. over what it would be if the first stimulus were bright. Hence, the results are in agreement with the predictions of a single channel model of the psychological refractory period.

FIGURE 1



RT to the second stimulus as a function of the intensity of the first stimulus. Each point represents the mean of 20 trials.

TABLE III

MEAN DIFFERENCE IN RT (IN MILLISEC.) TO A BRIGHT SECOND STIMULUS WHEN IT IS PRECEDED BY A DIM S_1 AND BY A BRIGHT S_1

Subject	*Mean difference in RT to the first stimulus at the two intensity levels*	*Mean difference in RT* to the bright stimulus when preceded by a bright or dim first stimulus*
1	34	30
2	31	32
3	26	20
4	42	36
	33	30

* RTs at the shortest ISI are omitted.
Each entry represents the mean difference on 20 trials.

C. *Comparison of two-choice RT and mean first RT*

The mean first RTs under all conditions are shown in Table II. A two-way analysis of variance was done for each intensity level of the first stimulus to determine the effect on the first RT both of the ISI between first and second stimuli and of the intensity level of the second stimulus. No significant differences were found in

first RTs over the eight ISIs employed. Similarly, the intensity of the second stimulus had no significant effect. No significant interaction effects were found between intensity level of S_2 and the ISI.

It is interesting to note, by comparing two-choice RT in Table I with mean first RT in Table II that the latter is approximately 44 millisec. longer than choice RT when only one stimulus was presented. In both cases a constant foreperiod of 1·4 sec. was employed. This increase is the same for both intensities of the first stimulus, and is approximately the same mean increase (42 millisec.) reported by Gottsdanker, Broadbent and Van Sant (1963) in a study designed to explore the difference between first RTs and RTs to a single stimulus.

To determine whether this delay was due to some inhibitory effect of the second stimulus or to a change in set when the subject knew that two stimuli would be presented, a comparison was made between mean first RTs when both stimuli were presented with mean first RTs on catch-trials where both stimuli were expected but only one was presented. The results are shown in Table IV. The difference

TABLE IV

COMPARISON OF CRT TO THE FIRST OF TWO STIMULI WITH CRT TO A SINGLE STIMULUS WHEN TWO STIMULI ARE EXPECTED

	CRT to a bright stimulus in millisec.		*CRT to a dim stimulus in millisec.*	
Subject	*First of two stimuli*	*Single stimulus (two expected)*	*First of two stimuli*	*Single stimulus (two expected)*
	$n = 20$	$n = 80$	$n = 20$	$n = 80$
1	250	251	284	285
2	277	278	308	329
3	275	274	301	307
4	265	262	307	303
Average	267	266	300	306

between the two RTs was not significant, suggesting that the observed increase in mean first RT was due, not to the inhibitory effect of a second stimulus, but rather to a difference in set when two stimuli are expected.

DISCUSSION

This experiment was concerned with examining the prediction of a single channel model of response selection that speeding up the afferent travelling time of the first stimulus by x millisec. would result in a decrease in RT_2 of x millisec. The results tend to support such a model.

One possible objection is that when the first stimulus was presented, subjects might have tended to orient more towards the dim stimulus than to the bright first stimulus, with the result that the second stimulus was more peripherally received, and hence more slowly conducted. This would have resulted in a relative increase in RT_2 after the dim stimulus. However, this possibility appears unlikely for several reasons. Firstly, the latency of a saccadic eye movement is in the range of 150–250 millisec. (Westheimer, 1954), so that except for the very largest intervals employed, no movement in the direction of the dim first stimulus would have

occurred before both stimuli had been presented. Although Zuber (1965) has recently found that some suppression of vision occurs prior to an eye movement, the latency of this suppression is about 50–80 millisec before the eye movement, so that at least for ISIs of 60 and 90 millisec., there should be no effect due to this suppression. Yet, with an ISI of 90 millisec., RT_2 is still faster following a bright S_1 than a dim S_1.

The extremely close correspondance between the decline in RT_1 with an increase in intensity and the decline in the latency of RT_2 following the more intense first stimulus suggests a high degree of independence in the processing of the two inputs, in a manner which allows almost a simple algebraic summation of components. If there had been some interaction between the two stimuli, such a clear relationship between intensity of S_1 and the latency of RT_2 would appear unlikely. It might also be noted that if the delay in RT_2 had resulted from a "refractoriness" in the system similar to that observed for nerves and muscles, as has been suggested by some authors (Telford, 1931; Davis, 1957), an increase in the intensity of S_1 might have been expected to cause greater refractoriness, and hence a greater delay in RT_2 following a bright S_1 than a dim S_1, rather than the reverse results which we found.

Concerning the effect of a second stimulus on RT to the first stimulus, the results indicate that the slight increase in RT_1 was the result of a change in set when two responses are required in rapid succession, rather than the actual presence of the second stimulus. This is supported by the finding that (*a*) RT_1 is not influenced by the interval between the two stimuli or by the intensity of S_2, and (*b*) the change in RT_1 is the same even if the second stimulus is not presented, provided the subject is expecting it.

To sum up, it appears that if the first response selection can be initiated x millisec. sooner, the latency of RT_2 is decreased by x millisec. Although this finding does not rule out various alternative explanations of the psychological refractory period, such as the expectancy theory (Elithorn and Lawrence, 1955) or the theory of competing responses (Reynolds, 1966), a failure to find it would have proved to be a serious difficulty for any single channel theory.

This investigation was supported, in part, by a fellowship granted by the Massachusetts Institute of Technology, and by the National Institute of Mental Health, grant MH 08890-02. The author is indebted to Wayne Wickelgren for his helpful suggestions.

References

Craik, K. W. J. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.*, **38**, 142–8.

Davis, R. (1956). The limits of the "psychological refractory period." *Quart. J. exp. Psychol.*, **8**, 24–38.

Davis, R. (1957). The human operator as a single channel information system. *Quart. J. exp. Psychol.*, **9**, 119–29.

Elithorn, A., and Lawrence, C. (1955). Central inhibition—some refractory observations. *Quart. J. exp. Psychol.*, **7**, 116–27.

Fraisse, P. (1957). La periode refractaire psychologique. *L'Année Psychol.*, **57**, 315–28.

Gottsdanker, R., Broadbent, L., and Van Sant, C. (1963). Reaction time to single and first signals. *J. exp. Psychol.*, **66**, 163–7.

Hick, W. E. (1948). The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36–44.

Reynolds, D. (1966). Time and event uncertainty in unisensory reaction time. *J. exp. Psychol.*, **71**, 286–93.

Smith, M. C. (1966). *Simultaneous vs. Successive Response Selection: A Study of Attention as the Limiting Mechanism.* Unpublished doctoral dissertation, M.I.T.

Telford, C. W. (1931). Refractory phase of voluntary and associative responses. *J. exp. Psychol.*, **14**, 1–35.

Vaughan, H. G., Jr., Gilden, L., and Costa, L. D. (1964). The functional relation of visual evoked response and reaction time to stimulus intensity. *Paper presented at Amer. Psychol. Ass. meeting.*

Welford, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance: A review and a theory. *Brit. J. Psychol.*, **43**, 2–19.

Westheimer, G. (1954). Mechanism of saccadic eye movements. *A.M.A. Arch. Opthal.*, **52**, 710–24.

Zuber, B. L. (1965). *Physiological Control of Eye Movements in Humans.* Unpublished doctoral dissertation, M.I.T.

Manuscript received 28th June, 1966.

SEQUENTIAL EFFECTS IN A TWO-CHOICE SERIAL REACTION TASK

BY

D. J. HALE*

From the Department of Psychology, University of Manchester

Twenty-four subjects performed a symbolic two-choice serial reaction task under four conditions. These were with a delay from previous response to onset of next stimulus of 100 millisec, 600 millisec., 2 sec., and a fourth condition of 2 sec. delay with verbal prediction of the next stimulus. A positive recency or repetition effect occurred at 100 millisec. delay where RTs to repeated stimuli were faster than RTs to alternate stimuli. At 600 millisec. this effect was still present, though much reduced. The 2 sec. delay gave a negative recency effect where RTs were slower to repeated than to alternate stimuli. This effect increased significantly with simultaneous prediction of the next stimulus. The verbal predictions themselves displayed negative recency. Run analysis of the four conditions revealed striking differences. These results emphasize the need for analysing the microstructure of choice RT situations and reveal deficiencies in present models.

Introduction

When one considers the extensive literature which exists on choice reaction time, it is surprising that an interest in sequential effects has not developed earlier. Sequential effects would easily be apparent in early serial RT experiments (such as those reviewed by Hansen, 1922) and more recently the effects were mentioned by Hyman in 1953, yet were ignored for seven years. It is also even more surprising that a sequential analysis was not earlier performed when we consider the close parallel between choice RT and probability learning situations where definite sequential effects were demonstrated by Jarvik as early as 1946. The importance of this sequential analysis of the microscructure of choice RT has recently been emphasized by Kornblum (personal communication), Williams (1966) and Falmagne (1965), who has attempted to account for frequency effects in terms of recency. Any model for choice reaction behaviour must explain the many sources of variation which affect results and explain the way in which specific signals are processed—not merely the results of an average process. The previous concentration on overall average performance has tended to obscure the importance of specific RTs to specific signals. For example, in a two-choice sequence, the individual RTs can be analysed separately into alternated or repeated responses, or a more detailed analysis can be made according to the position in a run of repeated responses. An effective model must be able to explain these detailed parts and not merely the overall RT to a sequence.

It is a frequent incidental observation in a discrete choice RT task that subjects tend to try and anticipate or guess what the next stimulus will be. When this guessing happens overtly, as in the case of Williams (1966), correct predictions give a shorter RT than incorrect predictions. The importance of this lies in the analogous situations of two-choice probability learning tasks. In these guessing situations (Jarvik, 1946, 1951; Nicks, 1959; Goodnow, 1955; Feldman, 1959; Winefield, 1966, and others) definite sequential effects occur. There is a tendency to predict alternations—as the length of a run of one event increases so will the tendency for the

*Now at Applied Psychology Centre, University of Aston in Birmingham

subject to predict the opposite event increase. Jarvik termed this effect "negative recency" and pointed out its most common manifestation in the well known "gambler's fallacy." If this tendency occurs in choice RT tasks it would be reasonable to expect that the bias of subjects towards predicting alternate events would lead to shorter RTs for alternate stimuli and longer RTs for repeated stimuli. This tendency was noted by Hyman (1953) for his two-choice condition, and also by Welford (1959) and Williams (1966).

Hyman (1953) also reported a positive recency effect in a situation with a greater number of alternatives, where RTs were faster to repeated than to alternate stimuli. In the first of a series of papers on the topic, Bertelson (1961), using a very compatible two-choice self paced serial task demonstrated a positive recency tendency which he termed a "repetition effect." He showed that under his experimental conditions the effect only occurred when the delay period from the release of the key from the previous response to the next stimulus was short—50 millisec. With a delay of 500 millisec. there was no effect; alternate and repeated RTs were equal. He suggested on the basis of his and Hyman's results that a repetition effect persisted longer when ". . . a more complex coding operation is involved. . ." (page 99). The more complex coding operation can consist of either a less compatible stimulus-response relationship or a larger number of alternatives. Both these variables were used by Bertelson (1963) who showed that the repetition effect increased as S–R compatibility was lowered by changing the spatial relationships of stimulus and response, and when the number of alternatives was increased. Bertelson and Renkin (1966) showed that in a less compatible task the repetition effect consistently decreased as delay periods were increased from 50 millisec. to 1,000 millisec. Leonard, Newman and Carpenter (1966) have also demonstrated a strong repetition effect in a highly biased five-choice serial reaction task.

The present experiment is designed to investigate the relationship between negative and positive recency and how they vary with the delay period in a serial task. The first objective is to determine to what extent positive and negative recency are dependent upon the delay period. It is predicted that positive recency will occur with short delays, that it will decrease as delay increases and will tend towards negative recency at still longer delays. The compatibility of the task used is intermediate between that used by Bertelson (1961) and that used by Bertelson and Renkin (1966) while the range of delay values is greater. The second objective is to assess the contribution towards negative recency or the attenuation of positive recency of the subject's guessing habits. This tentative explanation of the cause of negative recency was rejected by Williams (1966); however her experiment is subject to the criticism that she used verbal foreperiod warnings and thus may have given subjects cues as to which stimulus was to appear. The present experiment avoids this objection and also carries out the more useful sequential analysis of runs.

Method

Design

Each subject performed under four experimental conditions. A balanced design was used where 24 subjects were randomly assigned to the 24 possible permutations of orders of performing the four experimental conditions. The design was used to balance out possible intercondition effects which might have been particularly important because of lack of previous practice.

Apparatus

A self-paced serial choice reaction apparatus was constructed controlled by Hawker Siddeley LM10 logic modules. RT was measured by charging condensers from a stabilized 150 volt supply for a time equal to the duration of the S–R interval. The

resulting voltage was digitized and punched on to five hole paper tape by a Solartron Compact Data Logger. Serial measurement of RT to an accuracy of 5 millisec. was possible with basic delay periods as low as 50 millisec. by alternately charging, measuring and discharging a pair of independent condensers. The timing circuit was arranged so that 1 millivolt = 1 millisec. so that the condensers were charging on the linear portion of their charging curve. The basic delay could be varied by changing values of condensers which governed the time after keypress that the next stimulus appeared and the timer started. On a keypress the stimulus was extinguished and the next stimulus appeared a constant time interval after. The delays used were 100 millisec., 600 millisec. and 2 sec. Stimulus sequences were programmed on to paper tape and this tape advanced whenever either key was pressed—a non-correction procedure. An incorrect response caused an error to be recorded on a counter, as did simultaneous keypresses.

Stimulus-response details

Stimuli were either the number "1" or "2" displayed on a Mullard Z520M electronic number tube. Transparent perspex keys, resting on microswitches, were operated by left and right index fingers. They needed a pressure of 80 gm. to operate and emitted a distinctive click. The simplest S–R relationship was used with a "1" requiring a left keypress and a right keypress from a "2." Subjects sat comfortably in a screened cubicle with a low ambient illumination and free from distraction. They could hear the low noise of the tape punch after each keypress. The experimenter communicated to the subject through a loudspeaker; the subject, however, could not communicate with the experimenter during a sequence.

Procedure

Each subject attended for four separate half hour sessions. In each session he received sequences of stimuli of a given delay value. At the beginning of the first session each subject was familiarized with the workings of the apparatus and allowed a few practice responses under the three delay values. At the beginning of individual sessions subjects were shown the specific delay value for that session and performed a few practice responses. Apart from this there was no practice, except in the prediction condition, where 15 practice reactions and predictions were given to ensure that subjects could perform the predictions within the time required. With the 100 millisec. and 600 millisec. conditions, 20 sequences of 50 stimuli were given in four blocks of five sequences. All subjects performed the same 20 sequences in the same order. The 2 sec. and 2 sec. with prediction conditions were each of 10 sequences of 50 stimuli in two blocks of five, which were the first 10 sequences of the 20 sequences given in the shorter delay condition. There was a pause of about 20 sec. between each sequence when the subject was told the number of errors he had made on that sequence and the RT to the last stimulus. In the 2-min. interval between blocks of sequences the experimenter talked to the subject. In the verbal prediction condition the subject's predictions of the next stimulus were recorded on a tape recorder and transcribed later. The first RT of each sequence which was neither alternate nor repeated was removed in analysis.

Subjects

The 24 subjects were all members of Manchester University with a mean age of 23 years. Twenty were men and four were women. All were volunteers and received no payment.

Stimulus sequences

A large number of sequences of 50 random binary events (1 and 2) were generated from Fisher and Yates' random number tables. Sequences were selected for use which complied with the only constraint that the number of alternate events should be between 22 and 28 for either event. Table I shows the proportion of runs of various lengths for all 20 sequences and for the first 10 and last 10 sequences. They appear to be very similar in structure and the proportions of alternations and repetitions are equal in each case.

TABLE I

ANALYSIS OF STIMULUS SEQUENCES TO SHOW RUN STRUCTURE AND NUMBER OF ALTERNATIONS AND REPETITIONS

	Sequences of stimuli		
	1–20	*1–10*	*11–20*
Alternations	524	262	262
Repeat 2	253	136	117
Repeat 3	113	55	58
Repeat 4	56	26	30
Repeat 5	23	7	16
Repeat 6	8	3	5
Repeat 7	3	1	2

RESULTS

Errors

With the present apparatus errors could not be identified and removed, thus all average RTs include errors. The percentage errors were fairly low, however, with 100 millisec. 3 per cent., 600 millisec. 3·9 per cent., 2 sec. 1·8 per cent. and 2 sec with prediction 2·6 per cent. The low error rate and similarity between error rates for different conditions suggests that this will not be a source of error.

Alternate and repeated RTs

Table II shows mean RT and standard deviations for alternate and repeated RTs (and overall RT). By-subject analysis of variance was performed on the differences between alternate and repeated RT for all subjects and all conditions. This gave a between conditions F (3, 69) = 21·86 which is significant beyond the 0·001 level. A Neuman–Keuls test was performed on the average differences and showed that all are significantly different from each other beyond the 0·05 level, and that all but

TABLE II

MEAN RTs (IN MILLISEC.) AND STANDARD DEVIATIONS FOR OVERALL, ALTERNATE AND REPEATED RTs FOR ALL SUBJECTS UNDER THE FOUR EXPERIMENTAL CONDITIONS

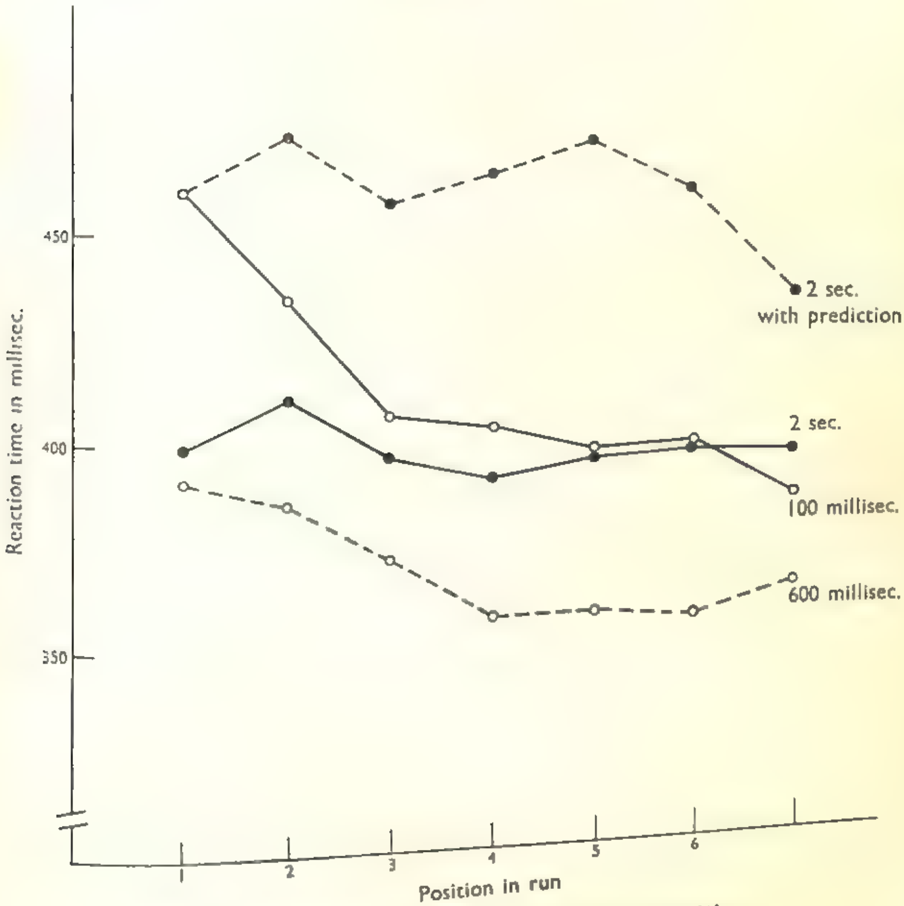
		100 millisec.	*600 millisec.*	*2 sec.*	*2 sec. with prediction*
Overall RT	Mean ..	441	385	401	464
	S.D. ..	106·7	80·4	72·3	149·5
Alternate RT	Mean ..	459	393	399	458
	S.D. ..	115·9	84·6	74·0	151·0
	N. ..	+12,000	+12,000	+6,000	+6,000
Repeated RT	Mean ..	420	376	405	470
	S.D. ..	93·9	76·0	70·7	148·0
	N. ..	+10,000	+10,000	+5,200	+5,200
Difference (Alternate – repeated)		+39	+17	–6	–12

the 2 sec. and 2 sec. with prediction conditions are significantly different at the 0·01 level.

Run analysis

The computer analysis gave mean RTs for stimuli in all run positions. The RTs for all run positions for each condition were superimposed, a procedure used by Bertelson (1961) and Leonard, Newman and Carpenter (1966), and the results are presented in Figure 1.

FIGURE 1

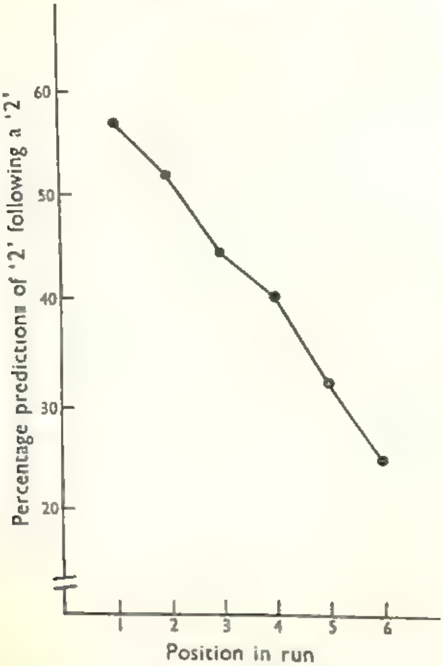


Mean RTs for position in run for the four conditions.

Analysis of predictions

After each occurrence of a stimulus of figure "2" on the number tube the subject would either predict another "2" or a "1". The prediction analysis is shown in Figure 2. This was constructed by noting the proportion of "2s" predicted after runs of different lengths of "2." These proportions were converted to percentage predictions. There are only runs of six in this situation because, as shown in Table I, the run of seven was for stimulus "1."

FIGURE 2



Prediction analysis for stimulus "2" showing percentage predictions of "2" for runs of "2."

Correct and incorrect predictions

Table III shows means for RT to correct and incorrect predictions for alternate and repeated stimuli. Analysis of variance showed that the only significant difference was between correct and incorrect predictions (F (1, 92) = 5·46) which is significant at less than the 0·025 level.

TABLE III

MEAN RTs (IN MILLISEC.) FOR CORRECT AND INCORRECT PREDICTIONS AS A FUNCTION OF ALTERNATE AND REPEATED STIMULI

	Alternate		*Repeated*	
	N	*Mean RT*	*N*	*Mean RT*
Correct predictions ..	2,636	423	2,481	422
Incorrect predictions	2,819	490	2,167	512

Discussion

The results clearly demonstrate the marked effects of variation in delay period on the overall reaction times and on the detailed structure of the reaction times making up these overall means. Comparison of these results with those of Bertelson (1961) and Bertelson and Renkin (1966) reveals a similar pattern with overall RT slower at the shortest delay (around 100 millisec.) than at intermediate values (around 500 millisec.). In the present study, lack of practice probably accounts

for the very large difference between the overall RTs for 100 millisec. and 600 millisec. This difference would be expected to decrease with practice, though, as Bertelson (1961) found, even after 8,000 reactions practice the shorter delay period still had the longer RT. The present experiment also confirms the general lengthening of RT with prediction reported by Williams (1966). The performance of some subjects in this condition was severely disrupted, but with others there was no effect.

The presence of a smaller repetition effect at 600 millisec. than at 100 millisec. confirms the general prediction of Bertelson (1961) that the repetition effect persists longer with the more complex coding operation, which in this case results from lower compatibility. The presence of this decreased repetition effect also confirms the results of Bertelson and Renkin (1966) who, however, only used delays up to 1 sec. and did not obtain a negative recency effect. Their patterns of alternate and repeated RTs have generally been confirmed, however, bearing in mind the lower amount of practice given in this experiment. Of particular interest in the present paper is the similarity of alternate RTs for the 600 millisec. and 2 sec. conditions while the repeated RTs display a large enough change to go from positive to negative recency. While Bertelson and Renkin did not achieve negative recency (presumably because of the delay values being shorter and the compatibility lower) they did report relatively steady alternate RTs from 200 millisec. to 1 sec. delays, while repeated RTs steadily lengthened. Thus over the delay values investigated the majority of changes in recency seems to be due to changes in the processing of repeated signals. The low value for negative recency at 2 sec. is probably due to the persistence of the repetition effect and competition between this and the causes of negative recency.

The demonstration that verbal prediction gives a significantly stronger negative recency effect differs from the results of Williams (1966) who reported no difference in recency, only a general lengthening in RT. This might be taken as a direct demonstration that negative recency can be caused by guessing strategies, and some support for this hypothesis comes from the result showing that the predictions show a strong negative recency tendency. The magnitude of this prediction effect is greater than previously reported (i.e. Jarvik, 1946 and Nicks, 1959) and this might be due to the highly paced nature of the task and the immediate high reinforcing value of making correct predictions. Correct predictions also lead to shorter reaction times, as was found by Williams; however, the non-significant recency effect seems to come mainly from incorrect predictions. This result, while differing from that of Williams who found that both correct and incorrect predictions gave almost equal amounts of negative recency, also causes difficulty for the guessing habits hypothesis. Reference to the run analysis curves raises another difficulty in relating the guessing strategies displayed to the RT recency effects. From the prediction curve one would predict that RT to successive repetitions would consistently increase. This is not the case and in the 2 sec. condition all the negative recency effect comes from the first repetition, with subsequent repetitions giving RTs very similar to the alternate RT. The prediction condition shows a similar pattern, with the larger negative recency effect being mainly caused by the first repetition. Subsequent repetitions again tend towards giving RTs equal to the alternate RT, though with greater variability. The status of the guessing habits hypothesis is then uncertain. The discrepancy between these results and those of Williams (1966) may be due to her use of verbal foreperiod signal, but the present experiment can also be criticized for the use of such a highly paced task and a situation which gives such a small negative recency effect.

Williams (1966) used a decision model to explain negative recency because of her rejection of its dependency upon guessing habits. While the status of the guessing

habits hypothesis is still uncertain, so is the status of her model. Williams seems to rather abruptly dismiss a repetition effect model, saying that a repetition effect reflects "... not properties of decision processing, but properties of a response system which has not fully come to rest between trials." If a repetition effect does not reflect properties of a decision processing system what then does it reflect? The major criticism of her model is that it makes no predictions for long runs of stimuli and certainly does not seem to be adaptable to the data for the 2 sec. and 2 sec. with prediction conditions presented here. Bertelson's (1963) model is also subject to similar difficulties in explaining the run curves for the 100 and 600 millisec. conditions. His model uses a classification system which asks a repeat question first, and then, in a two choice condition, will ask redundant questions if the answer to the repeat question is negative. This model does not explain the steady decrease in RT for the first four run positions in both 100 and 600 millisec. conditions, the different "terminal" values at which the runs level out or the fact that the repetition effect persists longer the lower the compatibility of the task. Bertelson assumes that a trace persists of the previous stimulus and decays with time, thus explaining why the repetition effect does decrease with time, as the repeat question cannot be asked reliably. However, it is necessary to add an additional assumption that the trace persists longer when the coding operations are more complex. The steady decrease in RT with successive repetitions to a terminal value could be explained by assuming that the repeat question is not always asked first, but that it tends to be asked first after being repeatedly reinforced by a series of repetitions. Thus the RT will decrease to some limit set by the responding mechanism when the repeat question is being reliably asked. Whether any reinforcement of this repeat question would have an effect on the number of redundant questions necessary to classify an alternate stimulus (and hence on alternate RT) could be a subject for further research.

It is interesting to speculate on the location or locations of the central processes responsible for repetition effects. Present evidence seems to point to some form of facilitation either involving the motor system or the later parts of translational activity. All previous experiments which have demonstrated repetition effects have used repeated responses with no intervening motor activity. If a motor or even muscular set could be preserved this might allow facilitation of a similar response. The effect of intervening activity was studied by Rabbitt and Rodgers (1965), who in a serial two-choice task made subjects give redundant responses to a "home" key after making a choice response. On touching the home key the next choice stimulus was presented and the whole cycle took 500 millisec., giving an effective delay equivalent to that of Bertelson (1961). It will be remembered that Bertelson had no recency effects at this delay while Rabbitt and Rodgers obtained a negative recency effect. Thus the redundant response seems to have forced a negative recency effect, which was, interestingly enough, only found in younger subjects. In effect, we cannot say that the negative recency effect was due to the breaking of a motor or perceptual set, but the fact that the response was redundant, requiring no choice, argues for a motor effect. This argument is supported by an ingenious experiment by Bertelson (1965) who varied stimulus uncertainty independently of response uncertainty. Subjects were required to make a common response to either of two stimuli and another common response to another two stimuli. The results were that only a slightly greater repetition effect was found when the identical stimulus was repeated rather than when the equivalent stimulus was repeated. The fact that there was a slight difference argues strongly for an analysis by sub-systems (such as in Welford, 1960) of the choice reaction process because there evidently is some perceptual component to a repetition effect.

I should like to express my thanks to Dr. Andrew Gregory for his generous help, and to Kenneth Pease and John Boddy for their advice.

References

BERTELSON, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quart. J. exp. Psychol.*, **13**, 90–102.

BERTELSON, P. (1963). S–R relationships and reaction times to new versus repeated signals in a serial task. *J. exp. Psychol.*, **65**, 478–84.

BERTELSON, P. (1965). Serial choice reaction-time as a function of response versus signal-and-response repetition. *Nature*, **206**, 217–8.

BERTELSON, P., and RENKIN, A. (1966). Reaction times to new versus repeated signals in a serial task as a function of response-signal time interval. *Acta Psychol.*, **25**, 132–6.

FALMAGNE, J. C. (1965). Stochastic models for choice-reaction times, with application to experimental results. *J. math. Psychol.*, **2**, 77–124.

FELDMAN, J. (1959). On the negative recency hypothesis in the prediction of a series of binary symbols. *Amer. J. Psychol.*, **72**, 597–9.

GOODNOW, J. J. (1955). Response sequences in a pair of two-choice probability situations. *Amer. J. Psychol.*, **68**, 624–30.

HANSEN, C. F. (1922). Serial action as a basic measure of motor capacity. *Psychol. Monogr.*, **31**, 320–82.

HYMAN, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188–96.

JARVIK, M. E. (1946). Probability discrimination and the gambler's fallacy in guessing. *Amer. Psychologist*, **1**, 453–4.

JARVIK, M. E. (1951). Probability learning and the negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, **41**, 291–7.

LEONARD, J. A., NEWMAN, R. C., and CARPENTER, A. (1966). On the handling of heavy bias in a self-paced task. *Quart. J. exp. Psychol.*, **18**, 130–41.

NICKS, D. C. (1959). Prediction of sequential two-choice decisions from event runs. *J. exp. Psychol.*, **57**, 105–14.

RABBITT, P. M. A., and RODGERS, M. (1965). Age and choice between responses in a self paced repetitive task. *Ergonomics*, **8**, 435–44.

WELFORD, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, **11**, 193–210.

WELFORD, A. T. (1960). The measurement of sensory-motor performance: a survey and reappraisal of 12 years progress. *Ergonomics*, **3**, 189–227.

WILLIAMS, J. A. (1966). Sequential effects in disjunctive reaction time: implications for decision models. *J. exp. Psychol.*, **71**, 665–72.

WINEFIELD, A. H. (1966). Negative recency and event independence. *Quart. J. exp. Psychol.*, **18**, 47–54.

Manuscript received 21st December, 1966.

OVERTRAINING AND GOAL APPROACH STRATEGIES IN DISCRIMINATION REVERSAL

BY

JEAN M. MANDLER and W. RONALD HOOPER

From the University of California, San Diego and Sir George Williams University, Montreal

An analysis of perseverative responding, position habits, and running speed in two experiments on Y-maze reversal learning is presented. These behaviours are described in terms of different goal approaching strategies in overtrained and mastery trained rats. Analysis of such strategies and other behaviour acquired during overtraining is applied to the problem of conflicting data and apparatus differences in the study of reversal and transfer learning.

Introduction

One of the most salient features of the large and growing literature on overtrained reversal and transfer shifts in the rat has been the inability to find stable replicable data. Theoretical proliferation has been the result; yet what is perhaps most needed is analysis of kinds of behaviour occurring in different apparatus, since apparatus differences appear to be at least partly responsible for the failures of replication. Two examples, out of many: D'Amato and Schiff (1965) failed to replicate Reid's Y-maze overlearning reversal effect (ORE) in their highly specialized and automated Y-maze, and Mandler (1966) found entirely different relationships between reversal and transfer in a Reid-type Y-maze than were obtained by Mackintosh (1962) using a jumping stand. The attempt to establish a general theory of reversal and transfer learning is hampered by the need to understand the peculiarities of various pieces of apparatus, each of which tends to emphasize different aspects of the discrimination process. Although several attempts have been made to transcend these difficulties theoretically by analysis in terms of attention and difficulty of discrimination (Lovejoy, 1966; Mackintosh, 1965), it seems unlikely that much can be accomplished in the direction of an adequate theory without more detailed information about the variables that affect choice behaviour in the various types of apparatus in which reversal and transfer data are collected.

Since theories based upon attentional variables tend to deal with problems of relevant and irrelevant stimulus dimensions (e.g. Mackintosh, 1965; D'Amato, 1965), it seems important to examine an apparatus in terms of the kinds of dimensions that it makes available to the subject and the kind of behaviour sequences that it encourages. This paper will present an analysis of several kinds of behaviour found in Y-maze reversal and then will relate these processes to behaviour patterns in other kinds of apparatus.

One of the most frequently reported types of behaviour in a two-choice visual discrimination reversal problem is perseverative responding; overtrained animals tend to choose the formerly correct stimulus in the early stages of reversal learning. Another frequent type of behaviour is position responding. Position responding has an anomalous status among the various "irrelevant" behaviours found in a visual discrimination problem. A position habit is not a stimulus dimension; yet it is frequently suggested as an irrelevant "hypothesis" or attempt at solution. It might be profitable to consider position habits not as attempts at problem solution but as

types of learned goal approaching sequences which both affect and are affected by the discrimination to be learned.

Such an approach was used in a previous paper (Mandler, 1966) on overtrained reversal and transfer learning in a correction Y-maze. Performance was analysed in terms of two goal approach strategies, one, a choicepoint strategy in which the animal consistently makes the goal entry choice in the choicepoint, and two, a detour or position strategy, in which the animal always enters the same arm of the Y and retraces on those trials in which the positive stimulus (S+) is in the other arm. During overtraining on visual discrimination problems, subjects gradually eliminated the use of the detour strategy and acquired a pure choicepoint strategy of goal approach. During transfer to another visual discrimination (extra-dimensional shift) the choicepoint strategy remained virtually intact and was associated with very rapid transfer. During reversal, however, the detour strategy was reacquired. There was some evidence of a negative relationship between the detour strategy (position habits) and reversal learning. At the same time, the presence of position habits was negatively correlated with perseveration to the former S+. As long as the animal continued to approach the former S+, it could not operate on the basis of a position habit. It was hypothesized, therefore, that increased perseveration to the former S+ by preventing the occurrence of interfering position habits might aid overtrained reversal learning and thus contribute to the overlearning reversal effect (ORE).

It was decided to test these relationships by examining other reversal data obtained from the same apparatus. Using the same Y-maze, Hooper (1967) investigated the effects of several variables on the ORE. His data showed that the ORE was more likely to occur under conditions of large reward than under small, with correction rather than non-correction procedures, and that distinctiveness of trials facilitated reversal but not differentially so between overtrained and mastery trained animals. His data, therefore, provided a wide variety of results on a reversal task which could be analysed in terms of degree of perseverative responding and the onset of the detour strategy. Running speeds in various parts of the maze during reversal are also reported.

Method

Procedure

Details of the apparatus and procedure are reported in Hooper (1967). In brief, the method was as follows: 128 hooded rats learned a black-white discrimination and its reversal. The variables in the 2 × 2 × 2 × 2 factorial design were: large reward (250 mg.) or small reward (37 mg.), correction or non-correction procedure, distinctive or non-distinctive trials, and overtraining or training to mastery. In the correction procedure the incorrect door was locked, and when the rat made an error it was allowed to retrace and enter the correct goal box. In the non-correction procedure both goal doors were unlocked but the animal could retrace at any point in the response sequence until it actually entered a goal box, at which point it was retained there. Distinctiveness of trials was achieved by removing the rat from the goal box between trials to a carrying cage, and by using a start box of different size, colour, and material from the goal box. Non-distinctive trials were achieved by retaining the rat in the goal box during the inter-trial interval and using that same box as the start box for the next trial. Subjects were given 10 trials a day. Mastery trained rats began the reversal task on the day after reaching a criterion of nine correct trials out of 10 for two successive days. Overtrained rats were given 200 additional trials before reversal. They were run 22 hr. hungry.

Apparatus

The apparatus consisted of a semi-automated Y-maze described elsewhere (Mandler, 1966). The start box had a guillotine door leading to the runway. Photocells and

timers registered elapsed times on Standard electric timing clocks; photocells were placed at the start box door, at the end of the straight portion of the runway, and just inside either arm of the Y. They measured the time elapsed since the preceding cells had been tripped. Timers at the goal box doors registered times at which the goal box doors were touched or opened by the animal. Stimulus cards, either black or white, were placed on the goal box doors and along the arms of the Y to the choice-point. On half of the trials the positive stimulus was in the left and on half in the right arm of the Y, in random order.

Treatment of data

The data for the correction groups only are reported here, for two reasons. One, no overlearning reversal effect was obtained in the non-correction, small reward groups. Two, there seems to be no satisfactory way of directly comparing degree of perseverative responding between groups run by the correction and non-correction methods. In the former case, the animal is eventually rewarded on every trial. During reversal it experiences the new positive stimulus (S+) on each trial even while continuing to run to the negative one (S−). The non-correction method prevents both experience with S+ and reinforcement during those trials the animal runs to S−. Various measures used to estimate perseverative tendencies for the correction groups were highly correlated, but for the non-correction groups the same measures were negatively correlated. The reason for this difference was that in the correction groups perseverative errors dropped out in a gradually decaying function. Non-correction animals showed very long strings of initial perseverative errors, but as soon as the first correct response was made, perseverative responding rapidly disappeared.

An analysis of trials taken to learn the reversal task, amount of perseverative responding, and onset of position responding was performed on eight groups of eight rats each in a 2 × 2 × 2 design, varying degree of training, size of reward, and distinctiveness of trials. Trials to reversal were measured by the number of days taken to reach a criterion of mastery consisting of two successive days of 90 per cent. correct responding. Perseveration to the former S+ was measured by the first errorless trial, i.e. the first trial on which the animal did not touch the incorrect S− (former S+) door.

Position habits were defined in terms of first alley entered. Although an animal operating with a position habit may make an error every time S+ is on the non-preferred side of the maze, this is not necessarily the case. At the time of mastery of the original discrimination, for example, many animals entered the same side of the maze every trial and retraced before touching the incorrect door. A position habit was scored when an animal made its first alley entry to the same side of the maze on at least eight out of the 10 trials per day. The first day on which an animal displayed a position habit was used to measure the onset of position responding during reversal.

Results

Table I summarizes the reversal results for the eight groups. It can be seen that in every case the overtrained group reached criterion faster than its mastery counterpart group, showed a greater degree of perseveration to the former S+, and later onset of position responding. Between group Spearman rank correlations ($N = 8$) were performed on these three measures. The correlation between degree of perseveration and trials to reversal was $-0{\cdot}768$ ($p < 0{\cdot}05$); between degree of perseveration and onset of position habits $0{\cdot}738$ ($p < 0{\cdot}05$); and between onset of position habits and trials to reversal $-0{\cdot}835$ ($p < 0{\cdot}01$).

At the bottom of Table I are added relevant data from the previous experiment by Mandler (1966). These were a small reward, non-distinctive, overtrained group and a small reward, non-distinctive mastery group, identical to Hooper's comparable groups in all respects except that the stimuli were presented only on the goal box doors and not extended to the choicepoint. When the two sets of data are combined, the correlations become: perseveration and trials to reversal, $r = -0{\cdot}900$; perseveration and onset of position habits, $r = 0{\cdot}773$; trials to reversal and onset of position habits, $r = -0{\cdot}776$ ($p < 0{\cdot}01$ in all cases).

TABLE I

	Mean day to reach criterion		*Mean first correct trial*		*Mean day of onset of position habits*	
	Over-trained	*Mastery*	*Over-trained*	*Mastery*	*Over-trained*	*Mastery*
Large distinctive	9·4	14·0	19·5	10·6	8·2	3·4
Large non-distinctive	10·5	15·4	21·6	10·0	6·5	4·9
Small distinctive	11·8	14·2	22·4	8·5	6·0	3·1
Small non-distinctive	13·8	15·4	10·4	9·5	4·4	2·5
*Small non-distinctive	17·9	19·1	8·0	5·0	4·4	2·0

* Data taken from Mandler (1966).

Analyses of variance were performed on the measures of perseveration, position habits, and trials to reversal for the eight groups. Degree of training was the only variable that was consistently a significant source of variance. Overtraining increased speed of reversal learning ($F = 23\cdot96$; $d.f. = 1\cdot56$; $p < 0\cdot01$), as did large reward ($F = 4\cdot46$; $d.f. = 1, 56$; $p < 0\cdot05$), and distinctiveness of trials ($F = 4\cdot09$; $d.f. = 1, 56$; $p < 0\cdot05$). There was an almost significant interaction between size of reward and degree of training ($F = 3\cdot72$; $d.f. = 1, 56$; $p = 0\cdot06$), suggesting that large reward affected overtrained performance somewhat more than it affected mastery performance and thus contributed to an ORE. In the analysis of position habits, overtraining was associated with later onset of position responding ($F = 17\cdot58$; $d.f. = 1, 56$; $p < 0\cdot01$), as was large reward ($F = 6\cdot81$; $d.f. = 1, 56$; $p < 0\cdot05$). Finally, in the perseveration analysis, only degree of training was a significant source of variance, with overtraining being associated with greater perseveration ($F = 13\cdot08$; $d.f. = 1, 56$; $p < 0\cdot01$). It had been expected that large reward would also be associated with greater perseveration but this variable did not approach significance ($F = 1\cdot27$; $d.f. = 1, 56$). Thus, the only variable that consistently affected trials to reversal, perseveration, and position habits, was degree of training.

Mean daily speeds were recorded in the following parts of the maze: speed in leaving the start box, runway speed, speed of making the first choice of alley entry (First Choice Speed), speed of making correct alley choice (Correct Choice Speed), and speed in entering the goal box. Correct Choice Speeds and First Choice Speeds on the original discrimination and the reversal for combined overtraining and mastery groups are shown in Figures 1 and 2 respectively.

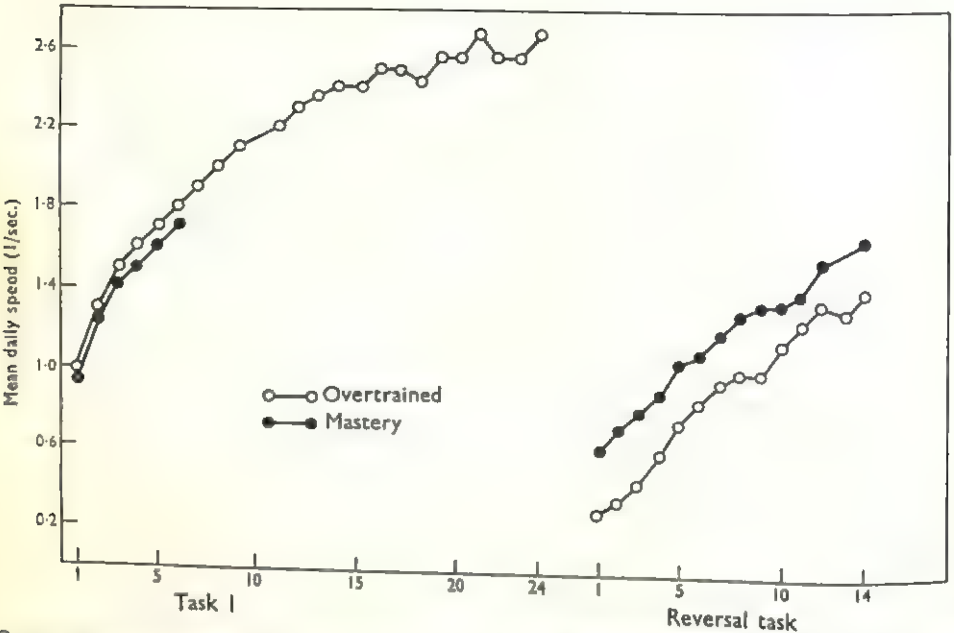
Analyses of variance on speeds during reversal showed that both large reward and distinctiveness of trials increased speed significantly on all five speed measures. Degree of training was also a significant source of variance, but its effects varied in different parts of the maze. The pattern of speeds associated with overtraining is similar to that reported by Mandler (1966).

Overtraining increased speed in leaving the start box ($F = 6\cdot01$; $d.f. = 1, 56$; $p < 0\cdot05$). On speed in the runway there was an interaction between degree of training and size of reward ($F = 4\cdot82$; $d.f. = 1, 56$; $p < 0\cdot05$). Overtraining resulted in faster speeds in the runway only for the large reward groups.

Figure 1 shows that overtrained groups were slower in making the correct choice ($F = 6\cdot63$; $d.f. = 1, 56$; $p < 0\cdot05$). They were also slower in entering the goal box ($F = 5\cdot17$; $d.f. = 1, 56$; $p < 0\cdot05$). Thus, overtraining resulted in faster running

speeds during reversal in the early portion of the maze sequence and produced slower speeds in the portions of the sequence having to do with choice behaviour.

FIGURE 1



Correct Choice Speed. Mean daily speed (averaged over 10 trials) of entering correct alley for overtrained rats ($N = 32$) and mastery rats ($N = 32$) on original discrimination and reversal. Values on the abscissa indicate days of learning.

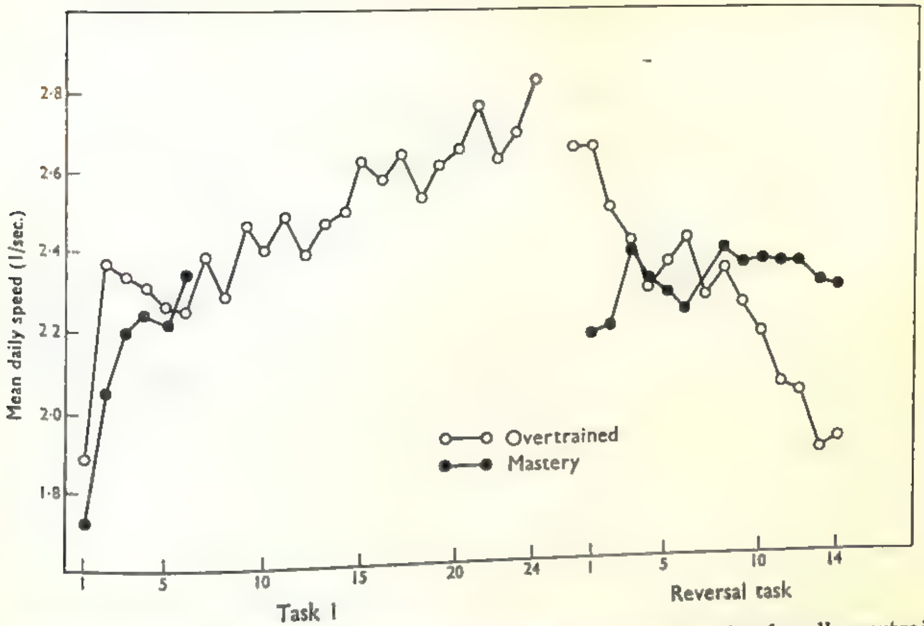
The First Choice Speed measure differed in this experiment from others using the same apparatus (see Mandler, 1966). In previous experiments stimuli were present only on the goal doors which were located 12 in. away from the choicepoint. In that situation speed in the choicepoint was an efficient predictor of position habits. During overtraining on a visual discrimination, speed of making the first choice of alley entry decreased in proportion to the elimination of position habits. In the present experiment the stimuli were extended along the walls of the alleys to the choicepoint, and speed was no longer correlated with position habits. In Hooper's data position habits were gradually eliminated during overtraining on task 1, but Figure 2 shows speed increasing during the same period. With stimuli immediately available it was not necessary for the animal to pause in the choicepoint.

During reversal, behaviour in the choicepoint was markedly different for overtrained and mastery animals. Figure 2 shows that overtrained animals, in spite of the proximity of the stimuli, slowed down in making their first choice as the reversal task was learned, while mastery animals remained at the same speeds attained during original learning ($F = 6{\cdot}60$; $d.f. = 13, 728$; $p < 0{\cdot}01$).

Discussion

The relationship between perseverative responding, delayed position habits, and rapid reversal learning previously noted in the Y-maze has been replicated in this study. It was suggested that perseveration is related to speed of reversal learning by virtue of its role in delaying the onset of interfering position habits. It is quite

FIGURE 2



First Choice Speed. Mean daily speed of making first alley choice for all overtrained rats and all mastery rats on original discrimination and reversal. Values on the abscissa indicate days of learning.

possible, however, that both perseverative tendencies and position habits are independent of each other and related to some other factor, such as attending to the relevant discriminanda. Mackintosh (1962) proposed that after overtraining, animals would continue to attend to the relevant stimuli and thus learn the reversal task faster, whereas mastery trained animals would abandon the relevant dimension during reversal and try out various other "hypotheses" before finally relearning to pay attention to the relevant dimension.

The present data are consistent with such a model of the ORE. Perseveration to the former S+ can be taken as direct evidence of paying attention to the relevant stimuli. It can also be argued that the presence of position habits in a visual discrimination indicates a lack of attention to the relevant discriminanda. Since overtrained animals show more perseveration and spend more time in the choicepoint, and mastery animals show more position responding and spend less time in the choicepoint, an explanation of the ORE in attentional terms would seem at hand. However, the relationship between position habits and attention is complex. Rats often consistently enter one side of the maze and at the same time pay attention to the discriminanda (or at least one of them) as evidenced by errorless performance. Nevertheless, the tendency to go to one side of the maze does increase the likelihood of errors on that side and thus implies some lack of attentiveness. This is especially true of mastery trained animals in the early stages of Y-maze reversal learning. The most typical pattern of errors is a brief spurt of responding to the former S+, followed by a consistent position habit with an error each time the new S+ is located in the opposite arm, and finally the gradual reduction in errors as the rat stops short of touching the incorrect door. This type of detour strategy seems to produce errors through sheer momentum; the rat bumps into the wrong door before it can stop

itself and turn around. It may also train the animal to pay more attention to S− than to S+, since it is only in the presence of S− that it must execute the fairly complex response of slowing down and turning around.

A choicepoint strategy avoids this source of errors since the choice behaviour takes place entirely within the choicepoint. It also tends to focus the rat's attention on both discriminanda at the same time. The latter advantage should show up primarily in shift or transfer learning, since continued perseveration during reversal tends to break down the choicepoint strategy.

The type of apparatus will influence the orientation of the animal to the stimuli and thus will influence the way in which the choice is made. Non-correction apparatus, such as a maze with doors to prevent retracing after the first choice is made, or a jumping stand in which the animal is retained on the incorrect platform, limit the range of choice behaviour open to the animal and prevent close contact with both S+ and S− on the same trial. This restriction presumably has an effect similar to that of the detour strategy in a correction Y-maze, namely, to emphasize the process of avoiding S−. In other words, apparatus differences which affect the pay-off for a correct or incorrect choice can be expected to affect the relative importance of S+ and S− in the learning process. Advantages and disadvantages for studying discrimination learning can be put forth for each of these experimental situations, but the point is that they not only emphasize different aspects of the discriminative process, they may also by their very nature affect variables influencing the discriminative process itself, such as the probability of attending to one or both stimuli.

One example will suffice to illustrate this point. In his study of reversal and nonreversal shifts in a jumping stand, Mackintosh (1962) found that overtraining increased speed of reversal, increased initial perseveration to the former S+, and decreased the number of position responses. These data are congruent with those presented here. However, on a shift from a black-white to a horizontal-vertical stripe discrimination, overtraining resulted in slower transfer learning and also in a greater number of position responses (although the latter difference did not reach significance). These results are directly opposite to those consistently found in our Y-maze, where overtrained transfer from a black-white to a stripe discrimination is very rapid and accompanied by virtually no position habits (Mandler, 1966). It seems unlikely that the differences found in the two apparatus could be accounted for in terms of difficulty of discrimination, since trials to criterion on the original task are not very different in the two studies.

The data suggest that the two apparatus lead to different ways of approaching discrimination problems. Overtraining in the Y-maze produces a well integrated sequence of responses leading up to and including the choicepoint, a sequence which fosters simultaneous scanning of the relevant stimuli. When faced with another task the animal continues the overtrained sequence. If the second task is a reversal, perseverative responding (which should be unusually persistent since the choice behaviour takes place in the middle of a well integrated sequence) eventually breaks down the previously learned sequence. If it is a visual transfer task there is nothing to prevent continued performance of the overtrained choicepoint sequence, with rapid transfer the result.

At the time of the second task, animals trained only to mastery (especially if the criterion is not a stringent one) have not yet developed a choicepoint strategy or any other consistent way of approaching the goal. The effect of the second task is to encourage the rat's original tendency toward a position habit. Whether or not a position habit is to be considered as an hypothesis, an attempt at solution, or a type

of goal approaching strategy, it is not a very efficient way of learning a visual discrimination. The animal receives information from only one stimulus at a time. The possibility also needs to be explored that this sequence produces more attention to S− than to S+.

A jumping stand seems to foster position responding. During reversal, perseverative responding by overtrained animals would have the same general effect as in a Y-maze, namely, to prevent position responding. Therefore, reversal data would be similar. Transfer data, in the absence of the old stimuli, could be expected to differ in the two situations. The return of position responding in the jumping stand suggests that the modes of scanning or other choice behaviour developed during overtraining is not of the same integrated stable character as that developed in the Y-maze.

The relationships discussed in this paper among the various behaviours learned during overtraining hold up under a wide range of conditions in Y-maze discrimination learning. The many replication failures, however, should alert us to the possibility that the ORE may be apparatus dependent. The similarity of trials to criterion and error data among various apparatus may have misled us into assuming apparatus differences are a minor factor in discrimination learning. This assumption is suspect on two grounds: First, trials or errors to criterion do not adequately represent what is learned; and second, a subsequent learning task will be affected by the various behaviours, including goal approaching strategies, that the animal has learned during the first task. A second task, therefore, is apt to be more sensitive to apparatus differences than origial acquisition.

Preparation of and the research reported in this paper were supported by Grant HD 01857 from the National Institutes of Health, U.S. Public Health Service.

References

D'Amato, M. R. (1965). The overlearning reversal effect in monkeys provided a salient irrelevant dimension. *Psychon. Sci.*, **3**, 21–2.

D'Amato, M. R., and Schiff, D. (1965). Overlearning and brightness discrimination reversal. *J. exp. Psychol.*, **69**, 375–81.

Hooper, W. R. (1967). Variables controlling the overlearning reversal effect. *J. exp. Psychol.*, in press.

Lovejoy, E. (1966). Analysis of the overlearning reversal effect. *Psychol. Rev.*, **73**, 87–103.

Mackintosh, N. J. (1962). The effects of overtraining on a reversal and a non-reversal shift. *J. comp. physiol. Psychol.*, **55**, 555–9.

Mackintosh, N. J. (1965). Overtraining, reversal, and extinction in rats and chicks. *J. comp. physiol. Psychol.*, **59**, 31–6.

Mandler, J. M. (1966). Behavior changes during overtraining and their effects on reversal and transfer. *Psychon. Monogr. Suppl.*, **1**, 187–202.

Reid, L. S. (1953). The development of noncontinuity behavior through continuity learning. *J. exp. Psychol.*, **46**, 107–12.

Manuscript received 5th September, 1966.

SHORTER ARTICLES AND NOTES

CONSERVATION OF QUANTITY IN CHILDREN: THE EFFECT OF VOCABULARY AND PARTICIPATION

BY

GILLIAN M. COHEN
From St. Anne's College, Oxford

In this pilot experiment some new tests for conservation of quantity were devised. The experimental group given these tests was able to make correct judgements, while a control group, matched for age and intelligence, failed to show conservation of quantity in the standard tests used by Piaget. The difference between the tests was analysed, and some factors emerged which, it is suggested, can serve to facilitate the child's performance.

Introduction

Piaget's work (1950, 1952) shows that children below 7–8 years of age lack the concept of conservation of quantity. Essentially they fail to realize that actual quantity can remain invariant in spite of perceived changes in the spatial relations. Further work has established additional factors governing the age at which conservation judgements can be elicited: for example, the mental age of the child (Feigenbaum, 1963); socio-economic status and length of schooling (Dodwell, 1960); the nature of the materials used (Elkind, 1961); and the complexity of the task (Feigenbaum, 1963). It has been suggested by Braine (Braine, 1964; Braine and Shanks, 1965) that younger children use and respond to words like "big," "more," and "same" in terms of phenomenal size rather than actual quantity. Such a tendency would obviously militate against giving the correct response in tests for conservation of quantity in which these words occur. It was therefore the aim of the present experiment to devise a new test which would circumvent a bias of this kind. This test has two features. (*a*) The "size" words ("bigger," "more," etc.) were eliminated from both instructions and responses. (*b*) The child was not asked to pass a verbal judgement on an existing state of affairs, but to perform a task bringing about a new state of affairs. The scores of the experimental group given this new type of test were compared with those of a control group given the standard Piagetian tests.

Method

Subjects. Two groups of 10 children each were used; a control group, and an experimental group matched for age and intelligence. Exact IQ scores were not obtained, but teacher's ratings of "below average," "average" and "above average" were used instead. The children were all at nursery school; they had a similar socio-economic background and approximately the same length of schooling. They were children of academic parents and formed an unusually intelligent group for their age. No attempt was made to match the groups for sex, as the sex factor has not been found significant in previous studies. Ages ranged from 4:2 to 5:9. The range was limited deliberately because it was anticipated that younger children would show no conservation under any circumstances, while older ones might well show it under both procedures.

Materials. In tests A and A′ and B and B′ two identical glasses referred to as X_1 and X_2 were used. The dimensions were—height = 3¼ in. and diameter = 1¼ in. A third glass, Y, was of height = 4¾ in. and diameter = 2½ in. For tests A and A′ small coloured sweets (Smarties) were used. For tests B and B′, orangeade. For tests C and C′, coloured plasticine.

Procedure. Tests and instructions were strictly standardized; the order of the tests was randomized for each subject; and no information was given about the correctness of his responses. This precaution was considered especially necessary since previous work shows that tasks involving discontinuous quantities (beads, counters) are found easier than tasks involving continuous quantities (clay, liquids), and transfer between tests has been found to occur (Smedslund, 1961). Each child in the experimental group performed tests A′, B′ and C′. Each child in the control group performed tests A, B and C.

Test A. The procedure is as described by Piaget. The child was shown the two identical glasses X_1 and X_2. He filled these with 25 sweets in each. The child took up a sweet in each hand, and dropped them simultaneously into the two glasses. He continued

to do so until the experimenter stopped him. The glasses then had 25 sweets in each and were two-thirds full. In case the child did not fully understand, he was assured that this procedure gave equal quantities in each glass. After he had agreed that they contained the same amount, the experimenter transferred the contents of glass X_2 to the empty glass Y (of different dimensions). X_1 and Y were placed together, and the child was asked "Does one have more or are they both the same?" In order to minimize factors of suggestibility, and limited attention, the form of this question was randomly varied in all the control tests. The alternative form was "Are they both the same or does one have more?" After his response, the child was asked "Why do you think so?"

Test A′. The experimenter showed the child the glass X_1 containing 25 sweets. He told the child "We are going to pretend that there are two children, Mary and Tommy, and we have to share out some sweets between them. We have to be absolutely fair, or else there'll be a terrible quarrel! Now these (indicating X_1) are Mary's sweets, and I want you to put Tommy's share in this glass (putting the empty glass Y beside X_1) and be sure you make it absolutely fair." The child could not proceed by counting the sweets in X_1 since not all were visible, and was therefore forced to make a spatial, not a numerical judgement. He was given plenty of time for the task and could revise his estimate until he was satisfied. He had unlimited supplies of sweets to work with. After he had finished, he was asked "How did you judge it?" If he had attempted to match the contents of the glasses height for height, his response was scored as non-conservation. If, on the other hand, he had clearly allowed for the greater width of Y, his response was scored as conservation.

Test B. The child was shown the glasses X_1 and X_2 filled to the same height with orangeade. He agreed they were the same. The experimenter transferred the contents of X_2 to the empty glass Y. X_1 and Y were then placed together, and the child was asked the same questions as in Test A.

Test B′. Instructions were as in A′. The experimenter showed the child the glass X_1 containing orangeade (the same quantity as in Test B). and told him "This is Mary's share." He gave the child a small jug of orangeade and told him to pour into glass Y a fair share for Tommy. Again the child had ample supplies to work with. He was asked to explain his judgement as in A′. Scoring was also as in A′.

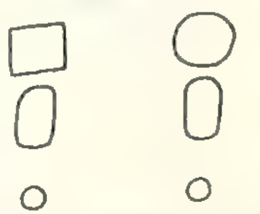
Test C. Equal balls of plasticine were shown to the child. He agreed they were equal. The experimenter then deformed one ball into a sausage shape. He then asked the child "Is one bigger or are they both the same?" as before, and why he thought so.

Test C′. The child was shown five pieces of plasticine of varying shapes—one sausage, two tiny balls, one round cake and one flat square. The balls were equal; the sausage was equal to the cake plus the square. The child was told the pieces represented chocolate.

FIGURE 1

The assignment of pieces planned for Test C′. The left hand group forms one share and the right hand group the other.

FIGURE 2



The assignment of pieces preferred by the children. The left hand side forms one share and the right hand side the other.

References

BRAINE, M. (1964). Development of a grasp of transitivity of length. *Child Develop.*, **35**, 799–809.

BRAINE, M. and SHANKS, B. L. (1965). Development of conservation of size. *J. verb. Learn. verb. Behav.*, **4**, 227–42.

DODWELL, P. C. (1960). Children's understanding of number concepts. *Canad. J. Psychol.*, **14**, 191–205.

ELKIND, D. (1961). The development of quantitative thinking. *J. genet., Psychol.*, **98**, 37–46.

FEIGENBAUM, K. (1963). Task complexity and IQ as variables in Piaget's problem of conservation. *Child Develop.*, **34**, 423–32.

PIAGET, J. (1950). *The Psychology of Intelligence.* New York: Harcourt.

PIAGET, J. (1952). *The Child's Concept of Number.* New York: Humanities.

SMEDSLUND, J. (1961). Acquisition of conservation of substance and weight in children. *J. Scand. Psychol.*, **2**, 71–84.

ZIMILES, H. (1963). A note on Piaget's concept of conservation. *Child Develop.*, **34**, 691–5.

Manuscript received 6th October, 1966.

VERBAL LABELLING AND LEARNING STRATEGIES IN NORMAL AND SEVERELY SUBNORMAL CHILDREN

BY

P. E. BRYANT

From the Medical Research Council, Institute of Psychiatry, Maudsley Hospital, London, S.E.5

The effects of verbal labelling of colours in a simple colour sorting task were examined in an experiment in which the subjects were severely subnormal and normal children of equivalent mental age levels. It was found that the effects of verbal labelling are relatively specific with severely subnormal subjects in comparison to their normal controls. It was also shown that this effect can be explained in terms of the relative independence with which the two responses in a sorting discrimination are learned by severely subnormal subjects both with and without verbalization.

Introduction

The aim of this experiment was to examine the effects of verbal instruction on the learning of a simple discrimination by normal and severely subnormal children. Previous studies (Bryant, 1964; 1965*a*; 1965*b*) of transfer of learning had provided evidence that the effects of verbal labelling are relatively specific with severely subnormal subjects. Overt verbal labelling in the initial task had improved the performance of the severely subnormal subjects in that task only, whereas it had improved the performance of the normal subjects in both the initial and the transfer tasks.

It was suggested that similar effects might also be found within a discrimination learning task, in which the subjects were made to verbalize about the relevant stimulus aspects. In order to examine the possibility that verbal labelling has relatively specific effects on the learning of a discrimination by severely subnormal subjects, an experiment was designed in which severely subnormal and normal subjects had to learn a two-choice sorting discrimination involving two colours. The two groups were divided into three subgroups, one of which received no verbal instruction, another verbal instruction (which had to be repeated) about both the colours, and the third instruction about only one of the colours. It was predicted that the main difference between the two groups would be in the pattern of the scores of the third subgroup. Verbal labelling of one colour would increase the number of correct placements by normals of both colours but would improve the number of correct placements by subnormals only of the colour which was labelled.

The pattern of learning scores predicted for the subnormals in the third group would indicate that the two responses were learned relatively independently by these subjects, since rapid learning of one response would not facilitate learning of the other. The question would then arise whether this relative independence in the learning of the two responses was a function of the introduction of the single verbal labelling or whether subnormals tend to learn the two responses independently both with and without the introduction of overt verbal labelling.

In order to test which of these two possible explanations was valid a post test was instituted, which was carried out immediately after the initial test. In this post test one of the two colours used in the initial test was held constant, and had to be placed in the same box, while a new colour was introduced, which had to be placed in the second box. No instruction was given in the post test. The rationale of this post test was that if the two responses in a simple sorting task are built up independently the existence of an established response would not facilitate learning of a new response.

Procedure

There were two groups, a severely subnormal and a normal group. The severely subnormal children attended a school in an institution for subnormals, and the normal children attended a local junior school. The two groups were equated in terms of mean M.A. of the subnormals and the mean C.A. of the normals. Both groups were subdivided into three subgroups, each consisting of 10 subjects. These subgroups were distinguished

Analyses of these results are presented in Table II. The analysis of the mean error scores for the three subgroups is given in Table IIA. This yielded a significant Groups × Subgroups interaction ($p < 0{\cdot}05$). t-Tests on the basis of this analysis showed no significant differences between the groups in the No-instruction subgroup ($t = 0{\cdot}460$, NS) and in the Double-instruction subgroup ($t = 0{\cdot}624$, NS), but did show a significant difference between the two groups in the Single-instruction subgroup ($t = 2{\cdot}765, p < 0{\cdot}02$). This showed that the subnormals made significantly more errors in this subgroup only. An analysis of the type of errors made by this subgroup is given in Table IIB. In this analysis the Response term refers to differences in errors made with the colour about which instruction was given and with the colour about which no instruction was given. The analysis yielded a significant Groups × Response interaction ($p < 0{\cdot}025$) showing that severely subnormal subjects made significantly more errors than normals in placing the colour about which no instruction was given.

Post test

In the post test one colour only was held constant from the initial task and a new one introduced. The only difference between the subgroups was that with the subjects in the Single-instruction group the colour held constant was always the one about which no instruction was given. With the subjects of the other two subgroups the colour held constant was systematically varied between subjects. The results of the post test are given in Table III.

TABLE III

MEAN ERRORS IN THE POST TEST

	Severely subnormals		*Normals*	
	No-instruction subgroup		*No-instruction subgroup*	
	Mean	*S.D.*	*Mean*	*S.D.*
Established response ..	0·8	0·98	0·7	1·02
Novel response	2·1	3·59	1·2	1·17
	Double-instruction subgroup		*Double-instruction subgroup*	
	Mean	*S.D.*	*Mean*	*S.D.*
Established response ..	0·5	0·68	0·4	0·66
Novel response	2·3	0·90	0·7	1·14
	Single-instruction subgroup		*Single-instruction subgroup*	
	Mean	*S.D.*	*Mean*	*S.D.*
Established response ..	0·8	0·89	1·8	1·40
Novel response	1·9	1·87	2·1	1·58

In Table III the results for the No-instruction and Double-instruction subgroups are given in terms of mean errors on the constant and novel stimuli. There appeared to be no difference between the subgroups within either group. Nor was there a difference between groups in terms of errors on the established response. However, in both subgroups the severely subnormal subjects made more errors in placing the novel stimulus than did the normals.

Similar results are given for the Single-instruction subgroup in Table III. This table shows that the severely subnormal subjects in this subgroup showed the same pattern as was found in the other two subnormal subgroups. They made more errors on the novel than on the established response. However, the scores of the normals in this subgroup were different from those found in the other two normal subgroups. In this subgroup

normals made more errors than the subnormals on both responses: this difference between groups was particularly evident with the errors on the established response. This result appeared to indicate that, when single instruction was given, normals are relatively handicapped in carrying over to a second task the response about which no instruction has been given. Since this pattern is different from that found with the No-instruction and Double-instruction subgroups, it can be suggested that, in the Single-instruction subgroup the normals, unlike the subnormals, paid relatively little attention to the non-instructed colour in the initial test, and in fact learned this task on an "*a*/not *a*" basis.

These post test results were subjected to analyses of variance, which are given in Table IV.

TABLE IV

Post Test Analyses

A

Analysis of the No-Instruction and Double-Instruction Subgroups' Scores

Source	*Mean square*	*d.f.*	*F*	*P*
Groups	9·112	1		
Subgroups	1·012	1	2·254	NS
Responses	19·012	1		
Groups × subgroups	0·613	1	1·365	
Groups × responses	6·613	1	14·728	0·001
Subgroups × responses	0·113	1	0·252	NS
Groups × subgroups × responses	0·612	1	1·363	NS
B.P.W.G.	57·15	36		
Residual	16·15	36		

B

Analysis of the Single-Instruction Subgroups' Scores

Source	*Mean square*	*d.f.*	*F*	*P*
Groups	3·600	1	0·870	NS
Responses	4·900	1	6·070	0·025
Groups × responses	1·600	1	1·985	NS
B.P.W.G.	74·500	18		
Residual	14·500	18		

In these analyses the Response term refers to differences between the established and the novel response in terms of errors made on these responses. The post test results of the No-instruction and Double-instruction subgroups were analysed together: the analysis is presented in Table IVA. This analysis showed a significant Groups × Response interaction ($p < 0{\cdot}001$), which was due to the fact that although the two groups made roughly the same amount of errors with the established response, subnormals made more errors on the novel response than did normals. The analysis of the post test results of the Single-instruction subgroup were analysed separately, since for these subjects the established response carried over was always the one about which no instruction was given. In this analysis, given in Table IVB, the only significant term was the Response term, showing that for both groups the novel response was the more difficult one. However, when the error scores made by the normal Single-instruction subgroup with the established response were t tested against equivalent error scores made by the normal No-instruction and Double-instruction subgroups, t was significant in both cases ($t = 2{\cdot}12$, $p < 0{\cdot}05$;

$t = 2{\cdot}43$, $p < 0{\cdot}05$, respectively). This showed that normals are relatively handicapped after single instruction in carrying over an established response to a new task, when this response is the one about which no instruction was previously given.

Discussion

Four main points should be made about the results of this experiment. The first is that both groups learn a simple colour sorting discrimination at the same rate, and benefit to a similar extent when both colours are verbally labelled.

The second point is that the effect of verbal labelling of only one of the colours is relatively specific with the subnormals in that they make more errors than normals in this condition, and this difference in error scores is confined to errors made in placing the colour which is not labelled.

The third point is that this relative independence in the learning of the two responses appears not to be a function of the introduction of the single instruction. The post test results indicate that when one colour is held constant and a new colour introduced, subnormals make considerably more errors with the novel than with the established response. Thus the subnormals, in contrast to the normals, appear to have to learn the novel response independently of the established response. Since no instruction was given during the post test, this is evidence that subnormals learn the two responses in a sorting discrimination relatively independently whether instruction is being given or not.

Finally, there was evidence that when single instruction is given, although normals learn both responses rapidly they are relatively handicapped in utilizing the response about which no instruction was given in a post test. This result suggests that the non-specific effects of single instruction in normal subjects are of a special kind, in that the two habits are learned on a "*a*/not *a*" basis. In other words, the non-instructed response is not learned on the basis of its stimulus properties but on the basis of it being the colour about which no instruction is given.

Hitherto work on the relationship of language to learning in children has concentrated on the qualitative changes in response strategies which are said to follow the introduction of verbalization (Luria, 1957) or of verbal mediating processes (Kendler and Kendler, 1962). These results, on the other hand, suggest that it is necessary to consider the effects of the introduction of verbalization in terms of the learning strategies which exist prior to any verbalization. The post test results provided evidence that the severely subnormal subjects learn the two responses in a sorting discrimination relatively independently, whereas normals do not. It is suggested that this difference in learning strategies is the basis for the finding that verbal labelling has more specific effects with severely subnormal than with normal subjects.

It has been reported that in comparison with normal children of equivalent mental age levels, severely subnormal children are relatively handicapped in the spontaneous formation of verbal connections (O'Connor and Hermelin, 1959; Bryant, 1965*c*, 1965*d*) and the use of these verbal connections in solving learning problems (Stevenson and Iscoe, 1955; Luria, 1963). The results of this study suggest that these handicaps may at least partly be explained in terms of learning strategies in severely subnormals and normals which exist independently of verbalization and which in subnormals are maladaptive to the introduction of verbalization.

The author would like to express his gratitude to the staff and children of St. Lawrence's Hospital and Dog Kennel Hill Junior School, where this research was carried out.

References

Bryant, P. E. (1964). The effect of a verbal instruction on transfer in normal and severely subnormal children. *J. ment. Defic. Res.*, **8**, 35–43.

Bryant, P. E. (1965*a*). The transfer of positive and negative learning by normal and severely subnormal children. *Brit. J. Psychol.*, **56**, 81–6.

Bryant, P. E. (1965*b*). The transfer of sorting concepts by moderately retarded children. *Amer. J. ment. Defic.*, **70**, 291–300.

Bryant, P. E. (1965*c*). The effects of verbal labelling on recall and recognition in severely subnormal and normal children. *J. ment. Defic. Res.*, **9**, 229–36.

Bryant, P. E. (1965*d*). The effects of verbal labelling on recognition of pictures and names in severely subnormal and normal children. *J. ment. Defic. Res.*, **9**, 237–44.

Gellerman, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *J. genet. Psychol.*, **42**, 206–8.

Kendler, H. H., and Kendler, T. S. (1962). Vertical and horizontal processes in problem solving. *Psychol. Rev.*, **69**, 1–16.

Luria, A. R. (1957). The role of language in the formation of temporary connections. In Simon, B. (Ed.), *Psychology in the Soviet Union.* London: Routledge and Kegan Paul.

Luria, A. R. (1963). *The Mentally Retarded Child.* Oxford: Pergamon Press.

O'Connor, N., and Hermelin, B. (1959). Discrimination and reversal learning in imbeciles. *J. abnorm. soc. Psychol.*, **59**, 409–13.

Stevenson, H. W., and Iscoe, I. (1955). Transposition in the feebleminded. *J. exp. Psychol.*, **49**, 11–5.

Manuscript received 5th September, 1966.

EXTINCTION OF A RUNWAY RESPONSE FOLLOWING RUNWAY OR GOAL BOX PARTIAL REINFORCEMENT

BY

R. WILTON*

From the Department of Psychology, University of Exeter

Rats were trained on a consistent reinforcement schedule in a straight runway. They were then switched to one of two partial reinforcement procedures. One group continued to run the full length of the runway, another was placed directly in the goal box. When extinguished in the full length of the runway both groups were more resistant to extinction than groups trained only on consistent reinforcement. An attempt was made to delineate the conditions for a demonstration of the partial reinforcement extinction effect. The results were discussed in relation to frustration theory.

Introduction

Amsel (1958) and Spence (1960) have introduced the aversive motivational construct "frustration" to account for differences in resistance to extinction following consistent reinforcement and partial reinforcement. It is assumed that a frustration response is elicited when an organism encounters an absence of reward in a situation where a reward has been previously present. For instance a frustration response is elicited on the non-rewarded trials of a partial reinforcement schedule. It is further assumed that elements of the frustration response condition to those stimuli which antedate it in a response sequence. In the typical runway situation antedating frustration responses could be elicited by stimuli prior to the goal area. The frustrative interpretation of the partial reinforcement extinction effect (PRE) is that instrumental approach responses condition to antedating frustration stimuli during training. Frustration stimuli then continue to elicit approach responses during extinction. Consistently reinforced animals, on the other hand, have no training to approach in the presence of frustration stimuli, hence the introduction of the aversive frustration stimulus in extinction results in a tendency to avoid the goal. Extinction thus proceeds at a comparatively rapid rate.

Frustration theory suggests that the conditioning of the instrumental approach response to the (antedating) frustration stimulus plays an important part in the determination of the PRE. The present experiment attempts to evaluate this suggestion by extinguishing a running response in four groups of rats. All four groups first ran a runway on a consistent reinforcement schedule. Two groups were then switched to a partial reinforcement schedule; one group continued to run the runway, the other was placed directly in the goal box of the runway. The extinction rates of these partially reinforced groups were then compared with each other, and with the extinction rates of the two groups which had remained on the consistent reinforcement schedule.

Method

Subjects. The subjects were 24 experimentally naïve male hooded rats, approximately 4 months old at the outset of experimentation.

Apparatus. The apparatus consisted of a runway, 6 ft. long, 4 in. wide, and 5 in. high. It had flat grey wooden sides, and a wire mesh floor and roof. The start box section, 12 in. long, was painted black. Latency and running times were combined and recorded on a timer. This was operated by a manually lifted guillotine door at the start box, and by a photocell situated at a point 2 in. from an aluminium tray, which opened into the end of the runway. A ½-in. diameter hole in the floor of the 1½ × 2 × 2½ in. tray enabled a rat to obtain a 50 per cent. concentrated sugar solution (50 gm. granulated sugar per 50 c.c. water) from a small fountain, produced just below the hole by a pump. The pump was activated by a break in the photobeam.

Pretraining. Each rat was individually handled for 5 min. per day on the first 3 days of pretraining. On the fourth and fifth days the rats were allowed to explore the runway

* Now at Laboratory of Experimental Psychology, University of Sussex.

for 5 min., the start door and sucrose pump having been removed. A 22¾ hr. deprivation schedule began on the fourth day, and for the remainder of the experiment each rat was given unlimited access to food for 1¼ hr. per day. On the sixth and seventh days the rats were allowed approximately 30 sec. access to the sucrose solution, at the reinforcement tray in the end of the runway.

Training. Training was divided into two stages.

Stage 1. Each rat was consistently rewarded in stage 1. A trial began with the placement of a rat in the start box, facing the door. Two-three sec. later the door was raised, starting the timer, and allowing the rat to run the runway. A break in the photobeam stopped the timer, and activated the sucrose pump. Each rat was given 30 sec. access to the sucrose solution, after which the pump automatically switched off. The rat was removed 2–3 sec. later.

Each rat was given a total of 50 rewarded trials; 10 trials per day for 5 days. The minimum inter-trial interval was 40 min. The rats were fed ¾ hr. after the last trial. Deprivation ranged from 16–22 hr. for each rat on each day of training.

Stage 2. The rats were arranged in order, according to their mean speeds on the last day of stage 1. This order was then split into six parts, and one rat from each part was assigned to one of four experimental groups. The four final groups thus showed approximately the same mean speed at the end of stage 1 training.

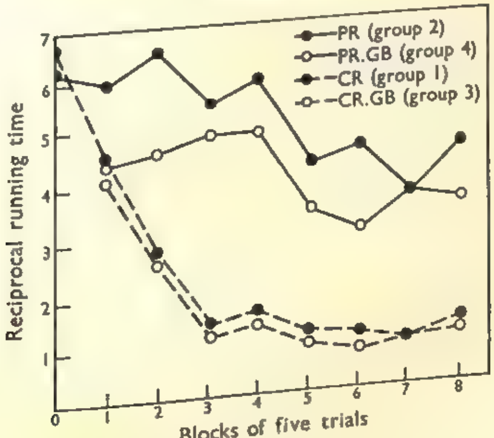
In stage 2 the four groups were treated differently. Groups 1 and 2 ran the full length of the runway, as before. Group 1 remained on a consistent reinforcement schedule. Group 2 was shifted to a 32 per cent. partial reinforcement schedule. For groups 3 and 4 the last 7 in. of the runway were sealed off by a block of wood, forming a goal box (GB). Each rat, on each trial, was placed in this goal box, *immediately in front of the reinforcement tray*. Group 3 was consistently reinforced on this procedure. Group 4 was reinforced on a 32 per cent. partial reinforcement schedule. Each rat was given a total of 50 trials, as in stage 1. On non-rewarded trials (groups 2 and 4) the sucrose pump (which made a slight hum) operated approximately 1 in. from the hole in the reinforcement tray, thus reproducing conditions found on rewarded trials.

Extinction. Conditions were similar to stage 1 training, except that the reinforcement pump was placed away from the tray, as in the partial reinforcement procedure of stage 2. A total of 40 extinction trials (4 days) was given. If on any trial a rat failed to break the photobeam within 90 sec. of the raising of the start door it was removed and accorded a running time of 90 sec.

Results

Training. In stage 1 mean speeds levelled off at about the thirtieth trial, remaining approximately constant thereafter. A Mann-Whitney statistical analysis showed that differences between the partially reinforced (group 2) and consistently reinforced (group 1) animals were non significant in the final trials of stage 2 (trials 45–50, $U = 9$).

FIGURE 1



Running speed during extinction.

Extinction. The mean reciprocal running time of each rat was taken over trials 25–40. Every rat trained on a partial reinforcement (PR) procedure showed a mean speed above that of every rat trained on a consistent reinforcement (CR) procedure. Thus a Mann-Whitney analysis ($U = 0$) gives a 0·002 level of significance (two tail) for differences between any PR and any CR group. A similar comparison of the two PR groups yields $U = 15$, and for the two CR groups, $U = 14$. Differences here do not approach significance.

It can be concluded that the partial reinforcement (PR) procedures are equally effective in producing a large partial reinforcement extinction effect (PRE) during extinction.

Discussion

The present results may be compared with those obtained by Trapold and Doren (1966). They used a procedure very similar to the present one. After preliminary training with consistent reward in a runway, several groups of rats were subjected to different partial reinforcement procedures. These involved either placing the rats in the goal box of the runway, with their noses directly in the food cup, or placing the rats in the goal box in such a way that they had to run 8 in. to the food cup. In both cases, of course, the food cup sometimes contained food, and sometimes did not. Some rats in both groups were also given interspersed consistently reinforced trials in the full length of the runway. In extinction the groups of rats fell into two clusters; the groups partially reinforced for running 8 in. to the food cup showed a PRE, whereas the groups placed directly over the food cup did not. The authors concluded that the PRE is response specific; that a PRE is demonstrated only when both rewards and non-rewards are administered contingently upon the response which is to be measured in extinction.

In the Trapold and Doren experiment the partial reinforcement of the running response seemed essential for the running response to show a PRE. But in the present experiment, the partial reinforcement of the running response did not seem essential, for it showed a PRE even though it had not been partially reinforced. The group placed at the reinforcement tray, in the present experiment, and the group which had to run 8 in. to the food cup, in Trapold and Doren's experiment, must have some training procedure in common (not the running response), which is not shared by Trapold and Doren's other groups which were placed with their noses directly in the food cup.

The crucial difference between those rats which show a PRE, and those which do not, seems to be that the rats *not* showing a PRE always have available in extinction, a set of stimuli which has been previously followed either by consistent reward, or by consistent non-reward.

Consider the partial reinforcement procedure whereby rats are placed in the goal box with their noses directly in the food cup. When a rat is placed over the food cup, signals are immediately available which are unambiguous with respect to the presence or absence of reinforcement. Food in the cup indicates consistent reward, absence of food indicates consistent non-reward. Hence when the rat is extinguished in the runway previously unambiguous signals are always available, those in the runway indicate consistent reward, those at the empty food cup indicate consistent non-reward. Rats run on this procedure do not show a PRE.

In the present experiment a partial reinforcement procedure was used which ensured that *only ambiguous* signals were available to the rat when it was positioned in front of the reinforcement tray. The rat was required to lick into the hole of the tray to determine the availability of reward. Unambiguous signals were available only when the rat's tongue made contact, or failed to make contact, with the sucrose solution. (Similarly if rats are placed in the goal box at a point 8 in. from a food cup only ambiguous signals are available between that point and the food cup.) When the rats were extinguished in the runway, previously unambiguous signals were not always available; signals in the runway indicated consistent reward, but signals at the reinforcement tray indicated intermittent (non-consistent) reward. Rats run on this procedure show a PRE.

It appears that the occurrence of a PRE depends on the significance that the stimuli, available at the termination of the response measured in extinction, have acquired during training.

Frustration theory (Amsel, 1958; Spence, 1960) assumes that the PRE is accounted for by the conditioning of instrumental approach responses to antedating frustration stimuli. However this assumption has little meaning until the operations are specified by which an approach response is identified. The phrases "locomotory response," "running the runway" and "approach responding" are often used as though they were

interchangeable. This is not so. The present experiment, for example, shows that a locomotory response can show a PRE even when no locomotory response has been partially reinforced, and when there has been no opportunity for one to condition to the frustration stimulus. The same cannot be said of an approach response, provided that the term "approach response" is defined in terms of its stimulus consequences, and refers to any action on the part of the rat which causes a specified set of stimuli (such as those associated with the reinforcement tray) to be presented to it. In the present experiment, stimuli from the reinforcement tray were produced by the rats during partial reinforcement training, and during extinction. This particular approach response (along with an infinite number of other approach responses) showed a PRE. If an approach response is appropriately defined then the present experiment does not demonstrate that a PRE can occur when no approach response showing the PRE has been partially reinforced. It does not therefore refute the hypothesis that the mechanism mediating the PRE is the conditioning of an approach response to the antedating frustration stimulus.

The author acknowledges the help of Dr. R. Brown who contributed greatly to the discussion.

References

Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychol. Bull.*, **55**, 102–19.

Spence, K. W. (1960). *Behavior Theory and Learning.* Englewood Cliffs, N.J.: Prentice-Hall.

Trapold, M. A., and Doren, D. G. (1966). Effect of noncontingent partial reinforcement on the resistance to extinction of a runway response. *J. exp. Psychol.*, **71**, 429–31.

Manuscript received 9th December, 1966.

FRUSTRATION, PREFERENCE AND BEHAVIOURAL CONTRAST

BY

T. M. BLOOMFIELD*

From the Psychological Laboratory, University of Exeter

Pigeons' pecks on both a red (left-hand) and a striped (right-hand) response key were reinforced on a concurrent variable-interval 2-min. schedule until the proportion of responses given to each key had stabilized. In alternate sessions, the right-hand key was covered, while responding to a green stimulus on the left-hand key was reinforced on variable-interval 1-min. When responding to green was later extinguished, more responses were made to the striped key in reinforcement sessions, although the rate of responding to the other, red key increased. Replacing extinction during green by reinforcement returned the preference and the response rates to their previous levels. These results are compared with a previous experiment in which the striped key was not present, where a similar increase in response rate to red was observed after extinction on green. The shift in preference coupled with the usual contrast effect in the present experiment supports an interpretation of behavioural contrast in terms of the frustrative effects of extinction.

Introduction

Behavioural contrast is an effect which appears in successive discrimination learning. In the usual procedure for training an operant discrimination (e.g. Hanson, 1959), a multiple reinforcement schedule is used. One stimulus (S+) is correlated with reinforcement, another stimulus (S−) with extinction, and S+ and S− periods alternate. Contrast occurs as an increase in the rate of responding to S+ over and above the rate to this stimulus before S− periods are introduced (Reynolds, 1961). The contrast effect shares some properties with the frustration effect obtained in the double runway (Amsel and Roussel, 1952). If reinforcement given customarily in the first goal box is omitted, then the running speed in the second runway is increased on that trial: in a similar manner, if periods of reinforcement for pecking alternate with periods of extinction, then the rate of pecking in reinforced periods rises. Another similarity is suggested in an experiment by Terrace (1963*c*), which showed that chlorpromazine reduced the contrast effect and disrupted a discrimination formed with many responses in extinction. An "errorless" discrimination (see Terrace, 1963*a*, *b*), however, in which no extinction responses had occurred, showed no contrast effect, and remained unaffected by the administration of the drug. Many experimenters (e.g. Barry and Miller, 1965; Boren, 1961; Dews and Morse, 1961) have reported the depressive effects of CPZ on the fear response, and some experiments on other depressive drugs show that frustration may be similarly reduced (Barry *et al.*, 1962; Stretch *et al.*, 1964). These results suggest the hypothesis that frustration is responsible for the contrast effect.

Amsel's frustration effect in the double runway is an immediate one: faster running occurs directly following nonreinforcement in the first goal box. Contrast, on the other hand, appears with long intervals separating extinction from the next reinforcement session (Bloomfield, 1967), indicating that it may be a function of conditioned rather than momentary frustration. Frustration responses in a multiple variable-interval extinction schedule (MULT VI EXT) may be conditioned to the stimulus present during extinction (S−), and hence generalize to the stimulus associated with reinforcement (S+). Responding in the presence of S+ could be increased by the motivational effects of generalized frustration, thus producing the contrast effect.

The experiment reported below attempts to test this hypothesis in a preference situation, where the pigeon has a choice between two response keys. In the ordinary one-key situation, frustration, according to the theory, increases response rate to S+ in MULT VI EXT. This has been demonstrated under the same stimulus conditions as those used below (Bloomfield, 1967). Because of the dual nature of the concept of frustration—both aversive *and* motivating—a specific prediction of performance in a

* Now at the Laboratory of Experimental Psychology, University of Sussex.

two-key situation can be made. If the second stimulus is not associated with S− along a generalization gradient, then this stimulus should be preferred to S+. The *rate* of responding to S+, however, should still increase with the presence of the EXT component of the multiple schedule, but the time spent on this key should diminish in favour of the second response key. Responding on this key escapes the aversive stimuli arising from generalized frustration.

METHOD

Subjects. Three locally obtained pigeons were used in this experiment. They were maintained at 80–85 per cent. of their free-feeding body weights throughout training. Fluctuations in body weights within these limits did not correlate with any phase of the experiment. All birds had previously performed in another discrimination learning experiment.

Apparatus. A standard experimental chamber (Ferster and Skinner, 1957) was used, with the exception that reinforcement consisted of one maple pea delivered by a modified pellet dispenser. At reinforcement the dispenser clicked audibly, the house lights switched off, and a small bulb over the pellet tray was illuminated for 2 sec. to allow the bird to eat. Stimuli were displayed behind each of two transparent response keys mounted at equal distances from the reinforcement tray. A measure of time spent in responding on each key in two-key sessions was taken by having a switch from one key to another start one timer and stop another timer. If the pigeon switched in the opposite direction, this operation was reversed. Other standard equipment controlled and recorded experimental events from an adjoining room.

Procedure. Since the three birds had undergone previous training in the same apparatus, no preliminary adaptation to the box or shaping of the pecking response was necessary. Reinforcement was made available on a variable-interval 2-min. schedule on each response key (CONC VI 2 VI 2) in a 1-hr. session. The left-hand stimulus was a red circle (S+), ¾-in. in diameter, and the right-hand stimulus a series of black slanting stripes on a white background (S_R). On alternate days, the right-hand key was covered, and the red circle was replaced by a green circle of the same dimensions. Responses to this one key were reinforced on VI 1-min., so that about 60 reinforcements per hour were delivered in both two- and one-key sessions. In phase I of the procedure, VI 1 and CONC VI 2 VI 2 alternated for 15 sessions each, 30 in all, during which the response rates appeared to have stabilized. In phase II, responses to the single green key were extinguished. CONC VI 2 VI 2 was still programmed on alternate days, and 10 sessions each of the concurrent schedule and extinction were run. Finally, phase III reverted to the conditions in phase I, and five more sessions of CONC VI 2 VI 2 alternated with five sessions of VI 1.

RESULTS

Figure 1 shows that the effect of the extinction of responses during green is to increase the proportion of responses per session given to S_R (stripes). When responding to the green stimulus is reinforced again in phase III, the preference returns to its previous level. The mean response *rates* (i.e. responses per key/time spent on the key) to S+ and S_R during the three phases of the experiment are shown in Table I. Response rates to S+

TABLE I

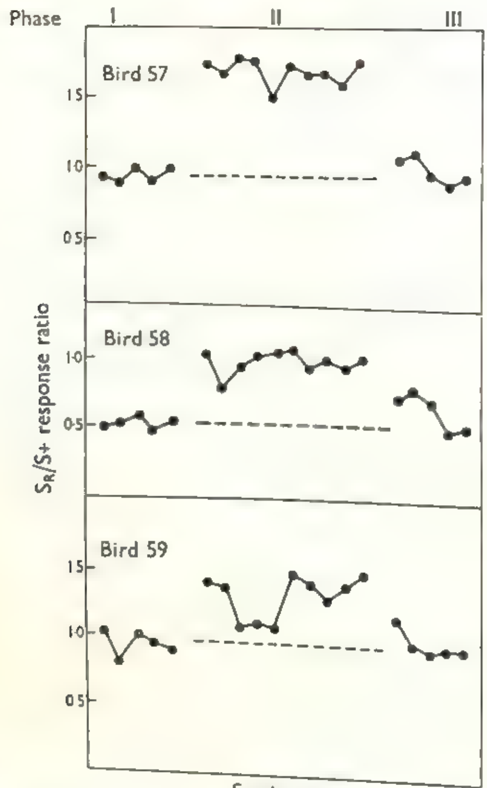
RESPONSES PER SECOND TO S_R AND S+, OBTAINED BY DIVIDING RESPONSES PER SESSION TO EACH STIMULUS BY THE TIME SPENT ON THE CORRESPONDING KEY

	S_R			S+		
Phase	*I*	*II*	*III*	*I*	*II*	*III*
Bird 57 ..	1·87	1·80	1·83	1·93	2·01	1·85
Bird 58 ..	2·05	2·00	2·10	1·91	2·30	1·95
Bird 59 ..	1·85	1·81	1·82	1·75	2·09	1·85

Figures are averaged over the last five sessions of Phases I and II, and the whole five sessions of Phase III. For measurement of time spent on each key, see under heading *Apparatus*.

show the usual contrast effect demonstrated before in the one-key situation, while S_R response rates show, if anything, a slight decline in phase II. In effect, the extinction procedure during phase II increases the rate of pecking on S+ when the bird is responding at this key; but it also increases the preference shown for the other key, S_R.

FIGURE 1



The ratio of S_R responses per session to S+ responses per session is shown for the last 5 days of phase I, the whole 10 days of phase II, and the whole 5 days of phase III. In phase II, responding to green (on alternate days) was extinguished. Results are given for three birds.

Discussion

The results provide clear support for a frustration interpretation of behavioural contrast. S_R (stripes) lies on a different stimulus dimension from S+ (red) and S− (green), and is on the right-hand side throughout training. So frustration induced by the extinction procedure and conditioned to the left-hand, green stimulus should generalize more strongly to S+ than to S_R. Thus when responses to S+ and S_R keys are reinforced equally often, the pigeon prefers S_R more in phase II than in phase I. When responses are reinforced in the presence of green again, in phase III, the preference reverts to its previous level.

The earlier study referred to above (Bloomfield, 1967) used the same procedure as in this experiment, with the exception that the S_R response key was absent. So responses to S+ and to S− were respectively reinforced and extinguished, on alternate days. Under these conditions, the usual behavioural contrast effect appeared. But although the increased responding characteristic of contrast occurs both in a one- and two-key situation, when a choice is presented, the stimulus preferred is S_R, not S+.

In this connection, it is worth noting briefly that there are analogies with this result. It is known that rats prefer consistent to partial reinforcement, although the running speed is often greater in 50 per cent. than in the 100 per cent. reinforcement conditions (Haggard,

1959; Wagner, 1961). Another similar result has been obtained by Fowler (1963) with electric shock. He ran two groups of rats, one rewarded in the goal box, the other shocked and rewarded. Some of the latter group ran faster than rats in the first group, although it seems unlikely that the shock condition would be preferred. These two results suggest that behaviour can be facilitated by aversive events. The experiment reported above supports the view that behavioural contrast can be treated along the same lines.

This research was conducted while the author was on a Science Research Council Studentship.

References

AMSEL, A., and ROUSSEL, J. (1952). Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, **43**, 363–8.

BARRY, H., and MILLER, N. E. (1965). Comparison of drug effects on approach, avoidance, and escape motivation. *J. comp. physiol. Psychol.*, **59**, 18–24.

BARRY, H., WAGNER, A. R., and MILLER, N. E. (1962). Effects of alcohol and amobarbital on performance inhibited by experimental extinction. *J. comp. physiol. Psychol.*, **55**, 464–8.

BOREN, J. J. (1961). Some effects of adiphenine, benactyzine, and chlorpromazine upon several operant behaviours. *Psychopharmacologia*, **2**, 416–24.

BLOOMFIELD, T. M. (1967). Some temporal properties of behavioural contrast. *J. exp. Anal. Behav.*, **10**, in press.

DEWS, P. B., and MORSE, W. H. (1961). Behavioural pharmacology. *Ann. Rev. Pharmacol.*, **1**, 145–77.

FERSTER, C. B., and SKINNER, B. F. (1957). *Schedules of Reinforcement*. New York: Appleton-Century-Crofts.

FOWLER, H. (1963). Facilitation and inhibition of performance by punishment: the effects of shock intensity and distribution of trials. *J. comp. physiol. Psychol.*, **56**, 531–8.

HAGGARD, D. F. (1959). Acquisition of a simple running response as a function of partial and continuous schedules of reinforcement. *Psychol. Rec.*, **9**, 11–8.

HANSON, H. M. (1959). Effects of discrimination on stimulus generalisation. *J. exp. Psychol.*, **58**, 321–34.

REYNOLDS, G. S. (1961). Behavioural contrast. *J. exp. Anal. Behav.*, **4**, 57–71.

STRETCH, R., HOUSTON, M., and JENKINS, A. (1964). Effects of amobarbital on extinction of an instrumental response in rats. *Nature*, **201**, 472–4.

TERRACE, H. S. (1963*a*). Discrimination learning with and without "errors." *J. exp. Anal. Behav.*, **6**, 1–27.

TERRACE, H. S. (1963*b*). Errorless transfer of a discrimination across two continua. *J. exp. Anal. Behav.*, **6**, 223–32.

TERRACE, H. S. (1963*c*). Errorless discrimination learning in the pigeon: effects of chlorpromazine and imipramine. *Science*, **340**, 318–9.

WAGNER, A. R. (1961). Effects of amount and percentage of reinforcement and number of acquisition trials on conditioning and extinction. *J. exp. Psychol.*, **62**, 234–42.

Manuscript received 14th November, 1966.

EFFECTS OF RESPONSE PACING ON CONDITIONED SUPPRESSION

BY

DEREK BLACKMAN

From the Department of Psychology, The Queen's University of Belfast

In this experiment, which employed a balanced design with two rat subjects, the frequency of reinforcement remained constant while the rate of operant responding was varied by means of a response pacing technique. At each of three response rates, 1-min. periods of noise were presented, and, as these periods ended, a slight unavoidable shock was delivered to the rat. This procedure resulted in suppression of the operant responding during the periods of noise. This behavioural change was measured by a suppression ratio, essentially a comparison of the response rates in the presence and absence of the noise. The suppression ratios varied in a systematic way during the experiment, denoting most conditioned suppression when the baseline rate of responding was high, and least suppression when this was low. It is therefore concluded that response rate is one factor determining the degree of conditioned suppression in this controlled experiment. The conclusion is corroborated by absolute measures of responding during the pre-shock periods of noise.

Introduction

The behavioural phenomenon known as conditioned suppression was first demonstrated by Estes and Skinner (1941). They found that rats' behaviour which was maintained by a schedule of positive reinforcement was suppressed during a stimulus which preceded an unavoidable shock, this shock being delivered independently of the rats' behaviour. Estes and Skinner attributed this behavioural disruption to conditioned anxiety, and suggested that this experimental procedure made it possible to quantify such anxiety. They attempted to do this by comparing the normal response rate of each subject with its rate of responding during the pre-shock stimulus. Some currrent workers conceptualize the phenomenon in a similar way (for example, Kamin, 1965); conditioned suppression is said to be the result of a classically conditioned emotional response (CER). Respondents originally elicited by shock become elicited during the pre-shock stimulus, because the requirements for Pavlovian conditioning are satisfied by the relationship between the pre-shock and shock stimuli. These respondents are incompatible with operant bar-pressing behaviour, and therefore the response rate falls during the pre-shock stimulus. Quantification of conditioned suppression is therefore thought to be an indirect measurement of the strength of the conditioned emotional response. Although this interpretation may be useful, only the term "conditioned suppression" will be used in this report, for it is concerned solely with the observable behavioural phenomenon identified by Estes and Skinner.

The amount of conditioned suppression is a function of the characteristics of the pre-shock and shock stimuli, but it is not only determined by such independent variables. Brady (1955) found that the conditioned suppression produced by a standard procedure extinguished more quickly with animals exposed to low-valued ratio schedules of reinforcement than with subjects responding on interval schedules. He suggested tentatively that these differences might be attributable to the differences in the rates of responding produced by these schedules. Thus the ratio behaviour was characterized by high rates of responding; Brady argued that this greater probability of responding led to a faster extinction of the conditioned suppression.

Lyon (1963) has found that an animal shows less conditioned suppression when the baseline behaviour receives a reinforcement on average once per min. (VI–1 schedule) than when reinforcements are obtained only once every 4 min. on average (VI–4). Lyon suggested, therefore, that reinforcement frequency was a determinant of conditioned suppression.

In both these experiments, adequate analysis of the determinants of conditioned suppression is confounded by a failure to control response rate and reinforcement frequency independently. Higher rates of responding were accompanied by higher frequencies of reinforcement in Brady's experiment, and in Lyon's study, the higher

frequencies of reinforcement generated higher rates of operant responding. Either theory can account for both sets of results. Lyon (1965) therefore attempted to rectify this by controlling reinforcement frequency while varying response rates. The degree of conditioned suppression did not change throughout this procedure, the birds failing to emit any operant response during the presentations of the pre-shock stimulus. Lyon therefore concluded that response rate is not a determinant of the degree of conditioned suppression. However, Blackman (1966) has claimed that this conclusion is not in fact established by the experimental results reported by Lyon. Blackman reported two experiments which prompted a contrary conclusion, for when reinforcement frequency was controlled, it was found that rats responding at high rates exhibited more conditioned suppression than animals responding at lower rates.

The experiment to be reported was designed to investigate this controversy further. While the frequency of reinforcements obtained by two rats was controlled, their rates of operant responding were changed. This was achieved by response pacing procedures (Ferster and Skinner, 1957), the rates being increased and decreased in a balanced design. Initially, the degree of conditioned suppression was only partial (i.e. the response rates during the pre-shock stimulus were lower than the control rates, but greater than zero); it was therefore possible to detect both increases and decreases in suppression.

METHOD

Subjects. Two naïve male albino rats were used. These were approximately 150 days old at the start of the experiment. They were maintained at 80 per cent. of their free-feeding weights by controlled feeding.

Apparatus. One test chamber (Skinner Box) was used. The experiment was automated by timers, relays and uniselector switches. Responses and reinforcements were tabulated by impulse counters and a print-out counter. A cumulative recorder provided a continuous record of the subjects' operant bar-pressing. The pre-shock stimulus was "white" noise delivered to a speaker in the test chamber. Shock was provided by a scrambled, pulsed source (Grason-Stadler E1064GS), and was delivered to the rats through the grid floor of the chamber.

Procedure. The rats were accustomed to the deprivation conditions and to the test chamber. They were then magazine trained, i.e. 45 mg. pellets of food were delivered to a hopper in the chamber at irregular intervals independently of the subject's behaviour. During two further sessions, one bar-press was necessary to obtain a pellet (continuous reinforcement—CRF). These sessions ended when each rat had received 100 reinforcements. Thereafter, the experiment may be conveniently divided into the following phases:—

Phase A (42 days): Each rat was exposed to one daily experimental session of 1 hr. duration. Reinforcements now became available to a bar-press at irregular intervals which averaged ½-min. (VI–½ schedule). It should be noted that when a reinforcement was made available by this programme, the tape-timer governing this did not "lock-up" in the conventional way (Stein, Sidman and Brady, 1958). Normally, when a tape-timer makes a reinforcement available to the next bar-press, it is stopped until that reinforcement is obtained by the subject. In the present experiment, the programme tape was driven continuously; as soon as a reinforcement was made available, this fact was stored by a relay until a bar-press occurred. Only one reinforcement could be stored in this way. Therefore a rat could miss a reinforcement completely if it failed to emit a response before the next reinforcement was made available. On the other hand, a rat could obtain 120 reinforcements per hour, even if it did not obtain each one immediately it became available. This arrangement was used throughout all phases of this experiment; the reasoning behind it will be discussed later.

On days 20 and 21 of phase A, adaptation to the neutral stimulus which would precede shock took place. "White" noise was presented for periods of 1 min. every 8 min. There were therefore seven presentations of the noise during each 1-hr. session. Counts were taken of the number of operant responses emitted during each 1-min. period of the noise, and these were divided by the number emitted in the immediately preceding minute. This provided a suppression ratio, a measure of any behavioural change associated with the noise. A suppression ratio of 1·00 is obtained if the noise has produced no measurable effect on behaviour, i.e. the rate of response during the noise is the same as the rate in the absence of noise. A suppression ratio greater than 1·00 represents an increase in response

rate during the noise, and a ratio smaller than 1·00 is produced by some degree of suppression during the periods of noise. Complete suppression is indicated by a ratio of 0.

During the last 20 sessions of phase A, conditioned suppression training took place. The independent variables were as for days 20 and 21, but as each period of noise ended, an unavoidable shock was delivered to the subject through the grid floor of the chamber. The shock generator was set to deliver a shock of 0·5 milliamps. intensity and 0·5 sec. duration. Suppression ratios were calculated as before.

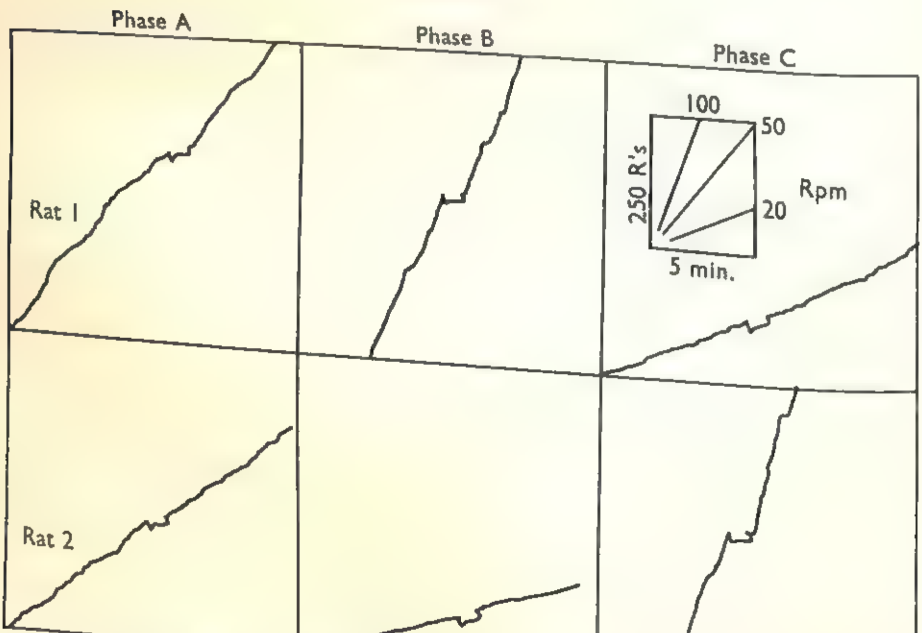
Phase B (30 days): The two rats were now exposed to paced schedules in which reinforcements continued to become available on average twice per min. (VI–½). However, for rat 1, only a response which occurred more than 0·05 sec. but less than 0·3 sec. after the previous response could be followed by reinforcement. This schedule may therefore be specified as VI–½ pacing (1) 0·05–0·3 sec. (see Ferster and Skinner, 1957, p. 498). Note, however, that there was no occasion in the present experiment when *every* response to satisfy the pacing requirements was reinforced, as was the case in the experiment reported by Ferster and Skinner. In this phase of the experiment, rat 2 was exposed to a schedule specified as VI–½ pacing (1) 5·0–8·0 sec., i.e. reinforcements became available at intervals varying about ½-min., but could be obtained only by a response which was emitted more than 5 sec. but less than 8 sec. after the immediately preceding response.

During the last 10 sessions of this phase, the conditioned suppression training was reintroduced with the same pre-shock and shock stimuli as were used in phase A of the experiment.

Phase C (30 days): The schedules were now interchanged, so that rat 1 was exposed to sessions of VI–½ pacing (1) 5·0–8·0 sec., and rat 2 to VI–½ pacing (1) 0·05–0·3 sec. During the last 10 sessions of this final phase, the conditioned suppression procedure was reinstated exactly as in the earlier phases.

Results

FIGURE 1



Segments of cumulative records illustrating the results. The pen is stepped across the moving paper by each response. Reinforcements are not shown on these records. The pen is deflected downwards during the periods of noise stimulus.

TABLE I

SUMMARY OF RESULTS

	Rat 1			*Rat 2*		
	S^d per min	*R.p.m.*	*SR*	*S^d per min.*	*R.p.m.*	*SR*
Phase A ..	1·96	46·4	0·43	1·97	40·1	0·28
Phase B ..	1·98	78·6	0·00	1·94	8·8	0·79
Phase C ..	1·91	12·9	0·72	1·98	90·6	0·05

Table I summarizes the results of the experiment. The figures recorded for reinforcements per min. (S^d per min.) and responses per min. (R.p.m.) are the means obtained from the last five sessions before conditioned suppression training began in each phase (before adaptation to the noise stimulus was begun in Phase A). Suppression ratios (SR) are the means obtained over the last five sessions of each phase, i.e. from 35 presentations of noise and shock. The data provided describe "steady states" of behaviour, for variations were slight and asystematic.

Segments of cumulative record are provided in Figure 1; these illustrate the results summarized in the Table.

In addition, the suppression ratios obtained during adaptation to the noise in phase A varied about 1·00. This reveals that the noise stimulus alone produced no measurable effect on the subjects' operant behaviour.

DISCUSSION

The simple pacing technique used in this experiment is certainly an effective method of varying response rate while the frequency of reinforcement is maintained at a fairly constant value. The two animals quickly adapted their behaviour so as to satisfy the timing requirements introduced in phases B and C. It should be added, however, that there was less stability within sessions in phase C. Rat 1, for example, made a small number of bursts of responding in the otherwise well-regulated behaviour. This might be interpreted as a result of the prior exposure to a schedule which reinforced such behaviour in phase B. The overall mean rates of responding obtained in phase C reveal that such irregularities were very occasional. The contingency whereby the tape-programmer governing the schedules did not "lock-up" in the conventional way when a reinforcement was made available helped to minimize changes in reinforcement frequency throughout the experiment. In a preliminary experiment similar to the present work, the tape-programmer was used in the normal way. It was then found that "lock-up" time accumulated throughout the experimental session, especially when the response rate was paced to a low value. Consequently, fewer reinforcements were available to the rat on this schedule, and a clear analysis of the results in terms of the effects of response rate was impossible.

In Table I it is clear that the suppression ratios obtained vary throughout the experiment. With both rats, the highest response rate was accompanied by most conditioned suppression, and the smallest degree of suppression occurred on the baseline provided by the lowest response rates. Sequence effects, which almost certainly occur, are counterbalanced by the design of the experiment. The very slight changes in reinforcement frequency cannot be thought to have affected the suppression ratios considerably. It therefore seems that response rate is the determinant of conditioned suppression emphasized by the results of this experiment. It should be noticed that suppression ratios are a *relative* measure of responding during the pre-shock stimulus, for this responding is compared with the rate of responding just before the noise stimulus is presented. It might be argued that the changes in suppression ratios in this experiment may be artefacts of this i.e. it might be the case that the absolute number of responses emitted during the pre-shock stimulus remained constant, in which case the changes induced in the control response rates would produce changes in the suppression ratios. The present results are not such an artefact, however. With both rats, the highest control response

rates were accompanied by the least amount of responding during the pre-shock stimulus in absolute terms, as well as by the lowest suppression ratios.

These results confirm the conclusion offered earlier (Blackman, 1966) that response rate is a determinant of conditioned suppression if the frequency of reinforcement is controlled. The repeated observation that high response rates are accompanied by most conditioned suppression may appear to contradict the results reported by Brady (1955), but it is crucial to realize that Brady failed to control reinforcement frequency in his experiment. This variable might also affect the degree of conditioned suppression, as Lyon (1963) has suggested. However, it is also clear that Lyon's proposal is affected by the present results, for Lyon varied reinforcement frequency but failed to control response rate. In fact, it seems that reinforcement frequency must act as a determinant of conditioned suppression in the way that Lyon suggests, for in his experiment the behaviour generated by the higher frequency of reinforcement was more resistant to suppression, *despite* the fact that it took the form of higher rates of responding which are now shown to be more easily disrupted by the procedure. Nevertheless, it would be prudent to confirm the previous findings concerning the effects of reinforcement frequency without the contaminations of the accompanying changes in response rates. It might be possible to do this by using response pacing techniques once again, this time to control response rates while reinforcement frequencies are varied.

Finally, only very slight changes in reinforcement frequency were produced by the large differences in response rates in the present experiment. This suggests that the response pacing procedure might profitably be used for a more precise analysis of the effects of response rate on conditioned suppression in which less extreme differences in rate are studied.

This experiment was indirectly supported by a grant from the Medical Research Council, for which the author is therefore indebted. The advice of Professor G. Seth in the preparation of this paper is also gratefully acknowledged.

References

Blackman, D. E. (1966). Response rate and conditioned suppression. *Psychol. Rep.*, **19**, 687–93.

Brady, J. V. (1955). Extinction of a conditioned "fear" response as a function of reinforcement schedules for competing behavior. *J. Psychol.*, **40**, 25–34.

Estes, W. K., and Skinner, B. F. (1941). Some quantitative properties of anxiety. *J. exp. Psychol.*, **29**, 390–400.

Ferster, C. B., and Skinner, B. F. (1957). *Schedules of Reinforcement.* New York: Appleton-Century Crofts.

Kamin, L. J. (1965). Temporal and intensity characteristics of the conditioned stimulus. In Prokasy, W. F. (Ed.): *Classical Conditioning: a Symposium.* New York: Appleton-Century Crofts.

Lyon, D. O. (1963). Frequency of reinforcement as a parameter of conditioned suppression. *J. exp. Anal. Behav.*, **6**, 95–8.

Lyon, D. O. (1965). A note on response rate and conditioned suppression. *Psychol. Rec.*, **15**, 441–4.

Stein, L., Sidman, M., and Brady, J. V. (1958). Some effects of two temporal variables on conditioned suppression. *J. exp. Anal. Behav.*, **1**, 153–62.

Manuscript received 9th January, 1967.

APPARATUS

AN AUDITORY AND VISUAL TIME MARKER FOR CLASSROOM USE

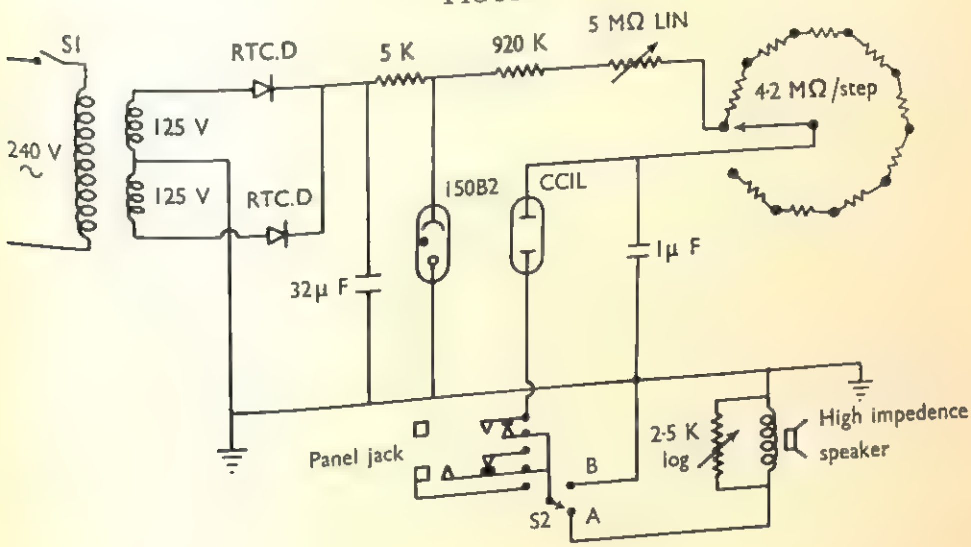
BY

P. J. COUGHLIN

From the Institute of Experimental Psychology, University of Oxford

Many situations used for classroom practical work require an easily manipulated time signal for presentation to either experimenter or subject. The traditional metronome suffers from the disadvantage that its volume cannot be controlled, with the result that it is difficult for more than one experimenter to work in a room at the same time. The "Auditory and Visual Time Marker" has been developed to overcome this difficulty. It generates visual and auditory signals at predetermined intervals (ranging from 0·25 to 10·00 sec.) which can be selected by setting a dial. The auditory signal may be presented through headphones or through a loudspeaker. In the latter case its intensity may be varied.

FIGURE 1



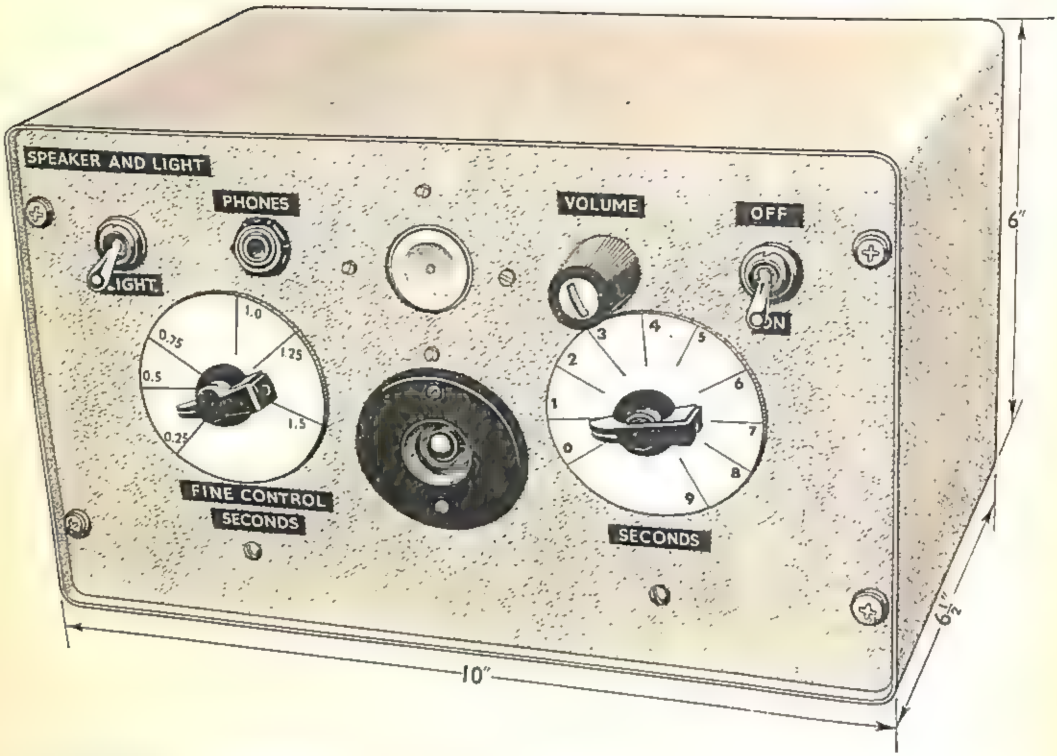
The circuit shown in Figure 1 is basically a relaxation neon oscillator with its own 150 v. stabilized power supply. The timing circuit consists of a fixed 1 μF. capacitor charged via a variable chain of resistance and discharged through a Hivac CCIL neon. The neon produces the light signal and the discharge current to earth via a loudspeaker produces the sound signal.

If the sound signal is presented through headphones, S_2 is switched to position B, thus isolating the internal speaker. The volume of the signal is constant and dependent on the impedance of the headphones used.

This instrument has proved very successful in practical class work involving time sampling and pacing.

The completed instrument shown in Figure 2 is relatively inexpensive and requires only minimal workshop facilities to produce.

FIGURE 2



My thanks to Professor R. C. Oldfield for permission to publish this article and to Professor E. R. F. W. Crossman for his encouragement.

Manuscript received 11th November, 1966.

BOOK REVIEWS

Patterns of Redundancy: A Psychological Study. By A. C. Staniland. London: Cambridge University Press. 1966. Pp. viii + 216. 45s. $8.50.

"*Patterns of Redundancy* is an attempt to develop the relationships between technical and less precisely defined uses of the word [redundancy] as they are relevant to problems of identification, discrimination, recall, and classification." The first half of the book is concerned with developing several measures of redundancy. Staniland claims that it is possible to measure redundancy objectively, with no consideration of an ensemble from which the actual sample is taken. He most closely considers psychological experiments, in which the sample consists of those stimuli actually displayed to the subject. One of the themes of the book is that the redundancy for such samples can be objectively measured without reference to the subject's ideas as to the ensemble from which the stimuli might have come.

Staniland sets up a great number of different measures, all giving answers in percentage redundancy. There are three main types of redundancy: L, S, and E. E is dismissed almost as soon as it is introduced; S, including O and $\sqrt{O}$, is discussed at some length, but no use for it is ever suggested; while L, including D and D_{min}, is used for any actual manipulations. The S series is based on the fact that the information in a redundant message could be transmitted with fewer different symbols in a message of the same length, while the L series is based on the fact that the information could be transmitted with a shorter message using the same symbols. He does not consider the idea that redundancy might be related to the comparison of the uncertainty of the message set with the uncertainty available using the same length and the same symbols. Computation from this (standard) notion of redundancy is numerically the same as using the L set, and in fact forms the basis for the D and D_{min} measures. Admitting the usual concept of redundancy would have halved the length of the book right away, since the S series would have been unnecessary, and the D series would be seen to result directly from variations in the implied ensemble.

Since Staniland does not admit that a measure of redundancy requires an implied ensemble, and devotes a good part of the book to attempts to show the contrary, he cannot admit the uncertainty ratio concept of redundancy as valid. Nevertheless, it becomes clearer with each reading of the book that redundancy in any of Staniland's measures depends crucially on the ensemble from which the stimulus sample is drawn, and more particularly in psychological experiments on the ensemble from which the subject thinks the sample might be drawn. D and D_{min} redundancy measures are merely devices to define apparently reasonable ensembles from which the stimuli actually presented might have been drawn.

It will be clear that I think the major thesis not proven. Failure to demonstrate the thesis is no reason to condemn a book. In fact, it might be a good reason to recommend it, if the thesis were well argued, and the book interesting to read. Controversial arguments are one of the ways a science progresses. This book, however, is written in an almost unintelligible manner, as witness, for example "It is a twist of language by which we tend to speak of the detail which contributes to specificity as itself specific" (p. 137), or "Behaviour according to a classificatory separation of redundant and required digits is a low level schematization. It issues in the strategy of concentrating the effort of specific storage upon only part of the display" (p. 136), or in discussing the presentation of the book (p. 9), "The two emphases which are particularly relevant are on the relational complexity of the assembly or source of displays, and on the part played by the participant agent as secondary only to the subject." I believe "participant agent" to mean experimenter.

The notation for measures of uncertainty is very difficult to follow. Four years after Garner's *Uncertainty and Structure as Psychological Concepts*, to which Staniland often refers, this is inexcusable. Perhaps it is due to the fact that this book was developed from a thesis presented in 1961, before Garner's was published. The book is also carelessly prepared, with explanations far from the place where the explanation is needed. Both points can be illustrated by the saga of "transmitted information." Garner uses $U(X:Y)$ to denote the contingent uncertainty, constraint, or information transmission between X and Y, and generalizes this to $U(X_1:X_2: \; : \; :X_n)$ for the total constraint among many variables. On p. 24, Staniland introduces $T(X;Y)$ for $U(X:Y)$, and $H(X,Y)$ for the joint uncertainty of XY pairs. He uses both symbols extensively on the following pages,

though on the next page the joint uncertainty has become H(X;Y). Thirty-seven pages later, by which time the transmission has become T(X : Y), and the total constraint has been introduced without explanation as $T_{1....c}$, one of the appearances of T(X:Y) triggers a footnote, explaining that he did not want to use Garner's notation because it "depends a great deal on differences in punctuation." By the time the book is finished, T has appeared in another guise, T(X,Y). One might almost sympathize with Staniland's comment, were it not for the fact that the chosen notation relies just as much as does Garner's on punctuation. I could go on at length about the difficulties introduced by the notation, style, and faulty argument, but lack of space forbids.

Apart from the unproven claims for objectivity in the measurement of redundancy, there is very little in this book that is not more carefully covered by Garner. My recommendation to prospective purchasers of *Patterns of Redundancy* is to buy *Uncertainty and Structure as Psychological Concepts* instead.

M. M. TAYLOR.

Signal Detection Theory and Psychophysics. By David M. Green and John A. Swets. London and New York: Wiley. 1966. Pp. xi + 455. 104s.

The need for a comprehensive introduction to signal detection theory has become increasingly strong in recent years, and a better pair of collaborators to satisfy this need could hardly be found. Both David Green and John Swets were trained at the University of Michigan, where signal detection theory was introduced into psychology by Wilson P. Tanner, Jr. In the past decade, these two authors have continued to contribute heavily to the technical literature on signal detection theory. But in writing this book, Green and Swets have done much more than satisfy a pre-existing and well defined need; they have in the book itself created a need which they then satisfy. This book leads as well as follows.

What is signal detection theory? In the broadest sense it is a theory of behaviour, because while it does not directly predict the outcome of most experiments in which it is used, it makes assumptions for its use which are clearly behavioural in nature. These assumptions are rather few but very important: A stimulus is assumed to lead to a sensory (or receptor or perceptual) process; the response made by a subject in a psychophysical experiment is a joint function of this sensory process and a decision process; the conceptually separate sensory and decision processes are capable of operational separation.

The operational delimitation of the sensory and decision processes requires that an additional form of data be collected in the usual psychophysical experiment: not only must we know the proportion of correct detections when a signal is in fact present, we must also know the proportion of incorrect detections when a signal is in fact absent. For a given sensory sensitivity, there is a whole set of points defined by these two values, points which differ due to some factors which cause a change in the subject's decision rule. Given such data, however, we can determine separately a value for the sensitivity and for the decision criterion.

A further assumption of signal detection theory is that decisions about sensory processes are made on the basis of a likelihood ratio, so that whenever a likelihood ratio can be generated, from whatever initial stimuli, signal detection theory is appropriate. The first of the three parts of this book is devoted to an explanation of these basic factors in signal detection theory, to comparisons with alternative theories (particularly with respect to the threshold concept), and to a demonstration of the validity of the signal detection concepts.

The authors take care to point out that signal detection theory exists in a psychological context. When the usual normal distribution assumption is made, signal detection theory leads to actual arithmetic operations which are in many cases identical to those found in Thurstonian scaling techniques. Furthermore, many of the more widely-used psychological procedures, such as two-category forced choice, are techniques which minimize the role of the decision process. Thus while these techniques do not provide separate measurement of the sensory and the decision processes, they do give a reasonably uncontaminated measure of the sensory process.

The measure of sensory sensitivity provided by signal detection theory places it squarely in the tradition of indirect measurement of sensory process, in that the measure does not assume a one-to-one correspondence between response and sensory process. While psychologists are far from agreed on the greater utility of the indirect measures, the increasing success of techniques based on signal detection theory makes clear that the

indirect techniques are most likely going to survive any onslaughts from more tradition-minded psychologists.

The weaknesses of this book are also its strengths, and three such paradoxical strength-weaknesses are worth some special mention. In Chapter 2 the authors discuss signal detection concepts in a strictly nonparametric fashion, without the use of the metricizing normal distribution. The metricizing assumption was actually made at the very beginning of the use of signal detection theory in psychology, so the present authors are creating a more general approach in first discussing it in a nonmetric fashion. But because it is so new an approach, there is little application of the more general approach to refer to, and therein lies the weakness of the very real strength of the more general approach.

Part II of the book is concerned with the concept of the ideal observer, especially in relation to auditory research. The concept of an ideal observer is a useful one and has been effectively used especially in relation to auditory work. The authors discuss this work thoroughly because there is application to relate to. However, this material will have much less general interest because it is so specialized in the sensory-energy problem areas.

The third paradoxical strength-weakness is illustrated in all of Part III, especially in the last chapter, where miscellaneous applications of signal detection theory are discussed. Here again the authors have too little material to allow the breadth of presentation that almost certainly will be available a few years from now. These new developments range all the way from animal psychophysics, to vigilance, to recognition memory, and its recent application to problems of short term memory is having a very substantial impact. So if the authors had waited a few more years to write this book, there would have been ever so much more to write. But if they had not written the book when they did, there probably would be much less about which to write.

How else can one summarize? These concepts and these techniques are here to stay, and will almost certainly be more used in the future than they are now. So this book must soon be used as a graduate text, for which it is well prepared with problems at the end of each chapter. But in the meantime, no human experimental psychologist can afford not to have a copy on a nearby shelf.

W. R. GARNER.

Acquisition of Skill. Edited by E. A. Bilodeau. New York and London: Academic Press. 1966. Pp. xiii + 539. £5.

This book consists of the papers presented at a meeting held in New Orleans during March 1965 with the support of a grant to Tulane University by the U.S. Army Medical Research and Development Command. It contains 10 main chapters each by a different author. All except the first and last are followed by a briefer contribution from another author whose task was to amplify and criticize the main statement. The first chapter, by A. L. Irion, outlines the history of research on the acquisition of skill and the last, by B. J. Underwood, attempts to tie together issues in motor and verbal learning. Between, C. E. Noble outlines a model to account for learning in multiple-choice situations; M. B. Jones addresses himself to the neglected problem of individual differences in learning and proposes a taxonomy of skills based on these; W. F. Battig discusses transfer effects; Ina Bilodeau surveys work on information about the results of action; the editor himself examines factors affecting retention and forgetting in motor tasks; and K. U. Smith discusses motor learning in cybernetic terms. Not all the chapters deal wholly or mainly with *acquisition* of skill: in particular E. C. Poulton, who surveys tracking studies, and J. A. Adams, who discusses observation and attention, are both concerned rather with *performance.*

The chapters will doubtless appear differently to different readers. The present reviewer found Poulton's and Smith's the most interesting. Poulton provides a remarkably lucid treatment of all the main aspects of tracking and gives unusually clear accounts of such difficult concepts as "display quickening" and optimum control gain. Smith summarizes some of the fascinating work he and his brother W. M. Smith have done on the spatial and temporal relationships between display and control, and especially on the effects of delayed or distorted sensory feedback on motor performance. Even if his arguments are, as F. Attneave who comments on his chapter calls them, "elusive" they are fresh and provocative. Reading these two chapters together one can scarcely fail to be struck by the immense potentialities of applying a small "on line" computer as used by the Smiths to tracking and other sensory-motor tasks. The detailed scoring and the measurement of parameters in transfer functions it would make possible, might

well render tracking tasks extremely sensitive indicators of fatigue and stress effects, and of impairment due to brain damage, age, drugs and other agents.

Each of the main chapters is a substantial piece of scholarship and is supported by a full list of references, so that the work is a mine of valuable information. This is the more so in that many of the references are to the unpublished reports of service and other research units: these reports are often difficult to obtain with the result that the work they describe attracts less attention than it deserves.

Having said this, however, it must be confessed that the overall impression of most of the chapters is disappointing. Except for Poulton's, Smith's, and to some extent Jones's they do not fully rise to the challenge of the theoretical problems posed by studies of complex skilled performance. Instead they merely use the results of these studies to further traditional approaches. The attitude is well expressed in the Preface:

> "The field of motor-skill acquisition is a part of experimental psychology and one of the siblings of the learning family. Those who work in the area think of themselves as contributing more to *learning* and less to *motor* and few investigators care much about motor-skill *qua* motor-skill."

The present reviewer profoundly disagrees with this restricted view, holding that research on motor skills during and since the war has played a far larger role than this, illuminating and developing many, perhaps most, areas of human experimental psychology. Those working in these areas today are often unaware of the ancestry of their ideas and methods and fail to realize the clearer understanding that would ensue if their origins, and the problems that gave rise to them, were recognized.

To put the matter another way, most of the chapters in this book lie on the extreme "right-wing" of present-day S–R psychology, dealing with overall trends of performance instead of examining it in detail, aiming at systematic description in terms such as "primacy-recency" rather than any more profound analysis or explanation, and seeking the "grand theory" rather than a simple understanding of behaviour. It is true that authors and discussants sometimes show they are aware of problems such as that a learning curve fails to convey a great deal of important information; that it does not say why, when a perfect performance has been attained once, it may not recur for many trials after; that errors may become ingrained and difficult to eliminate; or that skilled performances are flexible so that an expert cannot only do better what he has done before, but can do well what he has never done before. Little or no attempt is made, however, to consider questions of these kinds. The level of argument is strangely reminiscent of what was current in the Cambridge Laboratory 30 years ago. Indeed we were in some ways more advanced then because under the spur of Professor (as he was then) F. C. Bartlett and of G. C. Grindley we did at least ask questions like these urgently, even though we could scarcely then begin to answer them.

Such questioning seems to be fostered when the experimenter follows the better traditions of biological science in not being content merely to go through the motions of the hypothetico-deductive method but at the same time observes his subjects—and himself as a subject—in the manner of a natural historian aiming at an inductive insight into the phenomena he is studying. Without this observation, experimental psychology becomes an arid discipline indeed, breeding a jargon which can not only make reading extremely difficult but can distort ideas and facts by forcing them into, if one may mix a metaphor, preconceived strait jackets. This kind of jargon is all too evident in this book—Smith even seems to be *proud* of generating a new set of technical terms. Many of the chapters indeed seem almost to be worded as legal statements framed so as to make sure that nothing is left unsaid that ought to have been said, or said that ought not to have been said. As a result, the reader is often given tedious elementary definitions of tasks and concepts which he cannot possibly want if he is well enough versed in psychology to understand what follows.

Perhaps the most regrettable lack in the present volume is the absence of any sustained treatment of work on the acquisition of industrial or athletic skills. A good deal has been done in these fields, much of it, such as the work by Eunice Belbin and by E. R. F. W. Crossman and their associates, at a high level of theoretical sophistication. Yet these names are absent from the index, and other names associated with industrial training receive scant mention. Much of the impetus of present-day "pure" psychology came from the close association of theoretical and applied study during the 1940s and 1950s: it seems still vitally necessary that this association should continue if experimental studies are to maintain their perspective.

A. T. WELFORD.

Readings in Verbal Learning: Contemporary Theory and Research. Edited by Donald H. Kausler. London and New York: Wiley. 1966. Pp. xii + 578. 60s.

The readings in this book consist of 55 photocopied articles, of which almost half come from the *Journal of Experimental Psychology*, and nearly one-third from the *Journal of Verbal Learning and Verbal Behaviour*. They cover the years 1951–65, and the authors are all American. The book is intended for undergraduates with some knowledge of the subject and terminology.

Verbal learning is interpreted by Professor Kausler in a narrow sense, to include mainly rote learning studies and to exclude any work on reasoning, concept formation, grammar or syntax. Given this restriction the selected articles are fairly representative of current work; the only notable omission are studies of the different ways of assessing retention (total recall, recognition, etc.), and anything on mathematical approaches to the subject. The topics covered include both serial and paired-associate learning, serial position phenomena, isolation effects; the questions of one trial learning and of the "functional" stimulus; the importance of meaningfulness, familiarity, similarity, clustering, instructions to learn, and of mediating and interfering processes. The readings are classified under three headings, namely Acquisition (which receives the lion's share), Transfer, and Retention. Whilst one may feel that many of the articles assigned to Acquisition are also about retention little importance is attached to these distinctions. Kausler also provides a commentary interspersed between the readings; this constitutes about one fifth of the book.

It is difficult to see what justification there is for collecting together so expensively articles from such widely available journals. Although the selection is comprehensive in its range, and the standard of the articles mostly high, this does not seem the best way of promoting understanding of the subject. In the field of verbal learning there are few critical discoveries or dramatic innovations. Rather one is faced with a multitude of painstaking but, for the most part, individually unremarkable investigations; and any selection of a few articles is therefore unlikely to be very illuminating. What is more urgently needed is a critical account of the empirical generalizations established in the last 15 years, and of the implications of these.

The interspersed commentaries amplify the theoretical issues and mention relevant experiments not presented in full. There is too much repetition of what the authors themselves say. Furthermore the commentaries are in places extraordinarily difficult to follow, being written in a turgidly verbose style and liberally bestowed with symbolic abbreviations. The discussion is mostly judicious in tone, when not adulatory. Despite this, Kausler's attempt to defend the importance of verbal learning is extremely lame. Apart from simple assertions, with no specific examples, of the wide application of its concepts to other subjects, one of his main reasons (p. 3) is that psychologists in other fields, e.g. psychopathology, have found *confirmation* for their theories using rote learning techniques.

If verbal learning is to escape its somewhat cavalier dismissal as the dreariest chapter in psychology, it needs a livelier and more critical advocacy than this book provides. There are many interesting issues (e.g. the problem of defining the stimulus in serial lists), but they are too often lost in the surrounding babble of nonsense syllables.

JOANNA RYAN.

Experiments in Visual Perception. Edited by M. D. Vernon. London: Penguin Modern Psychology. Pp. 430. 8s. 6d.

This collection of readings in visual perception reprinted from the literature, should prove very popular. The articles are different from those collected by Beardslee and Wertheimer in 1958 (except for one), and 12 of them have been published since that date. Altogether it contains 35 papers, in seven sections.

The first four sections (269 pp.) cover experiments on the stimulus variables (form, space and distance, constancy and movement) which have traditionally attracted the greatest research effort. These range from "Old Testament" classics of the 1920s and 1930s, which will now perhaps be read as often as they are quoted—to Julesz's already classic work on binocular vision published in 1964.

The remaining three sections (161 pp.) cover experiments on personality variables (set and attention, motivation, and infant development) which have come into prominence recently. All these papers have been published in the last 20 years. They may not yet form a "New Testament," but we can already judge what may be gospel and what apocryphal. The picture has changed somewhat in the last 10 years, and what seemed

new and exciting a decade ago may appear now to have been premature. For example, Zubek *et al.*'s careful study on sensory deprivation (1961, reprinted here), indicates that the remarkable perceptual distortions reported by earlier workers may not be so easy to replicate as had been imagined.

The editor lets the different authors speak for themselves, with only 10 pages of introductory notes spread through the book to orient the reader. Instead of being embedded in exegesis, the articles themselves have been carefully chosen to lead the reader through a condensed history of the ideas on each topic. It is interesting to find the insightful demonstrations of Koffka and Gottschaldt juxtaposed with the quantitative, informationally oriented measures of Gestalt-like perceptual redundancy which were made by Leonard and Anderson, with references made to Attneave and to Rappaport. And the four papers on perceptual defence and subception (Bruner and Postman, 1947; McGinnies, 1949; Lazarus and McCleary, 1951; Rosen, 1954) may not convince the sceptical, but they do give a picture of the development of a line of research, as each experimenter fills in some of the gaps left by his predecessor.

The average undergraduate, or research worker, confronted by 30 years of perceptual research, is apt to be discouraged. This book will put new heart into him by making accessible a large number of articles which are usually inaccessible, or at least rarely sought out. It also offers a range of experimental methods and designs whose interesting little details are omitted by reviews and texts. By opening her files to us, Professor Vernon has made her own contribution to containing the information explosion. She captures only its brightest sparks in this far from deafening report from the frontiers of research.

The book is outstanding value for 8s. 6d.

S. M. Anstis.

Motivation. Edited by Dalbir Bindra and Jane Stewart. London: Penguin Modern Psychology. 1966. Pp. 352. 8s. 6d.

There is, at present, a spate of books of readings in Experimental Psychology; Penguin Books intend to contribute largely to this torrent, and this is one of their first titles. There is very little original editorial material in the book and nearly all the readings are fairly well known and easy to obtain; the aim of the book is, by the selection and organization of the material to illuminate the development of the concept of motivation. It is claimed by the publishers that the book provides "a coherent account of the development of theoretical ideas that guide current experimental work."

The readings are organized around three basic questions: The nature of drive, the problem of goal direction, and the relationship between drive and reinforcement. All but four of the 43 readings deal with the first and last of these topics; each of these longer sections is divided into sub-sections containing four or five readings on a subject such as "Drive as a Theoretical Construct" or "Reinforcers as Elicitors of Responses." This organization, together with the brief introductory notes to each section, gives this collection of readings the structure and individuality so often lacking in such compilations.

I am, however, doubtful of the value of this collection for two reasons. In those sections devoted to theoretical papers the excerpts given are often so short that it is hard to see how they can serve any useful purpose. For example Freud's views on instinct are given two pages and Lashley's get about the same; such brief passages cannot even give the barest outline of the complex systems that they are meant to illustrate. This enforced brevity of quotation also results in unacceptable bias; for example Hebb's important paper on "Drives and the C.N.S." clearly deserves inclusion, but it should not occupy as much of the book as the combined contributions of Freud, Lashley, Lorenz and Hull. The trouble is, not so much that the Hebb paper is too long to be included, but rather that far more space is needed so that equal justice can be done to other authors.

The second drawback to this collection arises in those sections of the book reprinting experimental papers. These are frequently given in full, and where they are cut this has been done with discrimination; in addition the papers themselves have been carefully selected. The objection is that there are too few papers in each subsection to give anything like a fair picture of the topic. For example the section headed "Drive Reduction as the Basis of Reinforcement," apart from about three pages of theory by Hull and Miller and Dollard, contains only two papers: Sheffield and Roby on the reward value of saccharine and Montgomery on exploration as a reward. Both these papers are, of course, highly relevant, but they can hardly be said to "provide a coherent account of the topic."

Both these criticisms are related. The readings selected are thoroughly appropriate,

but far too much has had to be left out, and so, through no fault of the editors, the book fails to achieve its declared aim. Had three books in this series been devoted to readings on motivation and had they all been entrusted to these editors I am sure that they would have succeeded. As it is the book must be regarded as a worthwhile attempt at an impossible task.

M. S. HALLIDAY.

Pattern Recognition: Theory, Experiment, Computer Simulations, and Dynamic Models of Form Perception and Discovery. Edited by Leonard Uhr. New York and London: Wiley. 1966. Pp. xii + 393. 68s. cloth, 45s. paper.

This book presents a selection of 22 papers on pattern recognition, with an introduction and linking passages by Uhr. The stated aim of the book is twofold: firstly to bring together material on pattern recognition from the different disciplines of psychology, neurophysiology, and computer simulation, and secondly to present a sample of computer models within this context. The book thus represents an attempt to provide a multidisciplinarian view of the problem of pattern recognition.

No one can deny that such an effort, if successful, is worthwhile, and in fact the study of pattern perception has already benefited greatly from the number of different approaches and techniques that have been applied to it. However, it is not clear that the selection of a small number of papers from the very large number available in the literature is a good way of helping to bring about such a synthesis. Some general principles could be applied, such as the selection of articles that are comprehensive, or controversial and stimulating, or difficult to get hold of, but within these limits the choice of papers has to be more or less arbitrary. It is difficult to find any consistent principle in the present case, and the linking passages do little to help in giving any overall view of the problem, so that the general effect is of a scrapbook of disconnected theories and experiments. The linking passages refer to other papers, on topics not covered by the published selection, which might have helped to increase the breadth of coverage provided by the book if references had been given for them. In spite of these limitations, however, most of the individual papers in themselves are interesting and stimulating, so that although the book does not fulfil its aims, it is nevertheless worth reading for this reason.

W. R. A. MUNTZ.

Advances in the Study of Behavior. Volume 1. Edited by D. S. Lehrman, R. A. Hinde and E. Shaw. New York and London: Academic Press. 1965. Pp. x + 320. 76s.

This book sets the reviewer a difficult task: it is a collection of papers, but not so numerous, bitty and overlapping as to give him a good excuse or incentive to make a general pronouncement, or take other evasive action. It is, on the contrary, a series of six serious and, on the whole, weighty treatises, none of which easily could be summarized in a sentence or two. The task is more or less that of reviewing six compact monographs, with relatively little overlap.

"Aspects of stimulation and organization in approach/withdrawal processes underlying vertebrate behavioural development" by T. C. Schneirla is, without a doubt, the weightiest of the lot. The paper is remarkable both for the breadth of knowledge and its opacity of expression. Schneirla elaborates his "biphasic" theory of approach and withdrawal at given stages of development for a variety of species, one of the main ideas being that development is based upon low-intensity stimuli (which are assumed to produce approach) reinforcing gradually more elaborate patterns of behaviour. Another key idea is that many of the preferences, say, in precocial birds, heretofore assumed to be stimulus specific by ethologists can be re-interpreted as being a non-specific function merely of sheer amount of stimulation, and that a heavy stimulus bombardment is aversive. In fact, Schneirla is proposing (as many others have done, e.g. Hebb, Duffy, much more explicitly) a kind of inverted U-shape curve for approach/aversion as a function of stimulus intensity. The argument is valuable and sometimes well-taken, but at other times it seems rather forced. For example, on the visual cliff the near surface "is seen as relatively unchanged whereas the deep one is seen as more in flux the greater the distance." Some of the material presented in this same volume by Walk can counter any such simple interpretation. My general opinion of Schneirla's paper is that it is provocative, opaque, over-inclusive, and valuable. I had to invent a new symbol for my marginal notes, which can best be described as a "quesclamation mark."

Prechtl's "Problem of behavioral studies in the newborn infant," somewhat surprisingly and even refreshingly does not concern itself with issues of nativism vs.

empiricism. It is, instead, a series of observations directed towards detecting individual differences and neurological disorders in infants. In fact, "behavioral" almost has the status in its title that "democratic" has in United Nations debates, the observations being concerned largely with general motor activity, posture, and the Moro response (a sudden clasping movement elicited by sudden change in body support or a sudden jolt). Prechtl has some quite interesting observations which will no doubt be of neurological value, but the paper cannot be described as being systematic or penetrating. It is fortunate his conclusion comes at the end of the paper, as I am not sure how many readers would stay the course if this were the beginning: "From the results of other investigators and our own research, the newborn infant is presented as a constantly changing creature, influenced by internal and external factors. Under the effects of these multiple influences, the baby's behavioral responses, in turn, are manifested in many patterns."

Of more direct interest to students of behaviour will be Walk's paper on "The study of visual depth and distance perception in animals." The author's work with Gibson on the visual cliff is well known and one can find here a clear and systematic review not only of that material but of other techniques, together with a discussion of the problems of dark-rearing and depth perception. Especially interesting are some of the historical tidbits and the variety of comparative evidence. The main approach is methodological and empirical, not speculative. While much of the material will already be familiar to the researcher on animal vision, the paper nevertheless will be a useful review to him as well as serving as a clear introduction to other readers.

The next chapter, "Physiological and psychological aspects of selective perception," by Horn, is a juicy combination of an admirably mustered review and judicious speculation. Horn first presents a sample of behavioural experiments from the general field of selective attention, and then turns to review the popular but complex field of electrophysiological counterparts of "attentive perception," habituation, and sensory interaction. The treatment, on the whole, is sophisticated; indeed, the criticism might be levied that it is too sophisticated. For example, Horn has a strong preference for single unit techniques, and is rather harsh about macro-electrode evoked potential work. "Studies of changes in amplitude of evoked potential have not led to substantial analytic advances in knowledge of the changes that take place in the sensory nervous pathways during attentive behavior. They have failed to provide unequivocal evidence on the question of whether or not the centrifugal system of fibers in the afferent pathways plays an important role" (p. 179). Similarly, he admits there is a good (although imperfect) relationship between brain rhythms and perceptual activity, but "it is not clear what changes in the transmission or processing of sensory signals the EEG changes express" (p. 182). These can be admitted, but there are many cases in physiological psychology where discoveries of correlations have advanced a field long before their underlying mechanisms have been understood. And there are instances where single unit recordings, despite their presumed superior methodological purity, correlate less well with behaviour than gross electrode recordings do (e.g. p. 182). Nor has the micro-electrode revealed those mechanisms as yet that the historically earlier findings failed to do. Horn's own single unit recordings of responses of tegmental cells to novel stimuli are important and clear. Although they unfortunately do not agree in detail with those from the Russian workers (e.g. Sokolov), it seems this line of work will certainly prove a fruitful one. It should perhaps be stressed that the physiological mechanism which does or does not provide for access to the tegmental cells is still quite unelucidated, and that still is the heart of the habituation problem.

The next paper, "Current problems in bird navigation," by Schmidt-Koenig, is almost orthogonal to all the others; it provides a wealth of detail about field and laboratory work of migration and homing, together with a review of recent hypotheses. It is a difficult chapter, but then so is the field itself (at least to this reviewer). It should prove a useful one even to the expert, since it is some years since a comprehensive review has appeared, but whether all experts will read it without emotional response is to be doubted; Schmidt-Koenig is not loath to levy diatribes. It is clear that the author is on the side of the angels in proposing a much more concerted application of statistical analysis to results and, in general, a tighter methodological control. The paper itself is unclear on a number of points, despite its strong organizational skeleton, and is rather overladen with minutia, but no doubt it will be much in use as a reference source.

The final paper is by Klopfer and Hailman on "Habitat selection in birds." It is a mixed bag (of the non-transparent variety) of rather narrowly based empirical data (the preference of Chipping Sparrows for pine foliage as compared with oak leaves) and what

the authors term "generalities and conjectures." The general problem of habitat selection appears to be one which is controlled by a complex intertwine of factors, and the treatment is probably by necessity heavily based on natural history. The observations are of some interest, but the paper has a much more hesitant and tentative character compared with the others in this volume. It is worth noting that the authors' own work seems to support the contention of Schneirla's (see above) that innate releasing mechanisms may be rather less specific than is usually thought.

The present volume is the first of an intended series which, according to the book jacket, "will be of great value to research workers and graduate students in endocrinology, psychiatry, neurology, physiology, experimental and clinical psychology, ethnology (sic?) and ecology." This first volume cannot be said to cover quite that scope, but it undoubtedly has something for almost everybody. Its content, generally, is of high quality and should prove durable. The title of the series is already self-congratulatory, but on the whole the results are to be praised. The volume itself is attractively presented.

L. WEISKRANTZ.

Fields of Psychology. Edited by J. P. Guilford. Third Edition. Princeton, N.J. and London: Van Nostrand. 1966. Pp. x + 350. 72s.

The first edition of this survey appeared in 1940 and the second 10 years later. The present edition had been extensively re-worked and contains several new chapters. It is virtually a new book.

Presumably, the aim of books such as this is to give students (or those considering becoming students) of psychology an idea of the subject in its major contemporary manifestations as well as the professional fields towards which its study may lead. If this is the objective, Dr. Guilford and his colleagues succeed admirably. If, however, one regards the function of texts to inspire as well as instruct, to instil critical sense and judgement, and to eschew the modish and merely parochial, this one hardly gets off the ground. It is conventional, pedestrian and in patches frankly boring. The best one can say for it is that it doesn't mislead and will do no harm.

O. L. ZANGWILL.

Attitudes. Edited by M. Jahoda and N. Warren. London: Penguin Modern Psychology. 1966. Pp. 375. 8s. 6d.

In an academic age that has been marked by the proliferation of "comprehensive" textbooks, it is a relief for a reviewer to have to deal with collections of readings. In the latter instance he is not required to assess the quality of the pap that has been predigested by the author for feeding to students—he is concerned with selection and omission. The present volume of readings shows careful selection, against a plan which includes various foci. We have an introduction, followed by some rather bitty pieces on the concept of attitude, then sections on research in attitudes (Content, Origins, Change, Behaviour), and, finally, Theory and Method. Two criticisms so far: the titles of papers are not given in the contents list at the beginning of the collection, merely the authors' names; excerpts from papers or chapters are often unsatisfactory in that one cannot follow the full development of the author's thought.

However, there are some excellent long papers cited in full. Particularly valuable is the Cook and Selltiz article reprinted from the *Psychological Bulletin*, 1964. And, to give the exception to the general rule above, the excerpt on "Attitude Scaling" by Selltiz, Jahoda, Deutsch and Cook is a most useful one. One of the central difficulties of work in this field has been in the establishment of outside criteria, whether behavioural or physiological, to show the presence or absence of a particular attitude or state of feeling, thus circling the pitfalls of verbal report. The papers in this volume pay a pleasing amount of attention to relationships between behaviour and attitudes, and physiological concomitants are not neglected—though more work in this area would have been welcome.

Omissions: It is a pity that the recent article by S. S. Stevens "A Matrix for the Social Consensus" (*Science*, 4th February, 1966) was too recent for inclusion, with its useful linking of psychophysical method and research into opinions and attitudes. There are two or three very slight references to Repertory Grid or Semantic Differential techniques in the text, and these popular procedures are nowhere explained. There are no technical references to language and linguistic research—crucially bound up with this field. Nothing from Thurstone—old fashioned?—certainly wise and a major pioneer in the field of scaling of attitudes.

But these are possibly carping criticisms—the selection is, overall, a good one, tending to provoke thought rather than to attempt definitive answers. The editors have

concerned themselves with attitudes in the areas of "modern" problems, and the book's influence should prove salutary in a society more and more manipulated from the centre and through mass media of communication. The present Penguin series, of which this is one, will be welcomed by many University teachers, as the provision of duplicate volumes of journals becomes ever more difficult. The demand for books on "Attitudes" is enormous, especially as the concept spans the discipline of psychology and the interests of sociology. Not even your reviewer's distrust of the concept of attitude could lead him to regard this as anything but a bargain, at 8s. 6d., for all undergraduates interested in social psychology.

S. G. Lee.

Personality Assessment. Edited by Boris Semeonoff. London: Penguin Modern Psychology. 1966. Pp. 443. 8s. 6d.

This book will be useful to the many who are concerned with the assessment of human personality, whether for purposes of research, teaching or clinical and vocational studies. It is unusual in that it does precisely what the editor claims for it, namely to include a wide variety of theories and practices dealing "with variations in 'normal' personality" and differing attempts to describe and classify these variations. The book consists of an Introduction by the editor; 19 chapters—each being a previously published paper or chapter—by an authority in the field of personality assessment and each preceded by a brief editorial note. At the end of the book is a short bibliography, an author index and a subject index. References are given at the beginning and end of every chapter.

In his Introduction, Dr. Semeonoff divides the chapters into five main groups although, as he says, "none of these groups represents an entirely homogeneous category, and there are certain overlaps between the groups as regards both orientation and method." The first group, which consists of five papers, is described as the most miscellaneous. It begins with an extract from Galton's *Inquiries into Human Faculty*, on Mental Imagery. This is characteristically wide-ranging, vivid and insightful. It is followed by Allport on Expressive Behaviour, Sheldon and Stevens on the Varieties of Temperament, Jung on Psychological Types and Hilgard on Experimental Approaches to Psychoanalysis. Thus this group might be designated as typological, philosophical and psycho-analytical in approach.

Semeonoff describes the second group as "operational" papers. "The approach in all of these is sometimes described as 'global,' in contrast to the 'dimensional' approach" which comes later. These papers concern the British and the American Officer Selection procedures, which were developed in World War II, a paper by MacKinnon entitled The Nature and Nurture of Creative Talent and a paper by Schafer on The Expression of Personality and Maladjustment in Intelligence Test Results. This group is perhaps the least impressive of the five: several chapters are depressing in their use of unnecessary jargon, e.g. "snafu tolerance" and their wordiness, e.g. "It may indeed be urged that the time has come for present-day psychologists, as a scientific group, to give, as their predecessors once gave, rather more attention to the status and developmental character of human knowledge and its relation to action, to the variety of methods through which knowledge is acquired, to the degree of certainty and precision which can be expected in a given area of knowledge at a particular historical moment, and to the relevance of such knowledge, certain and uncertain, precise and imprecise, to the solution of practical problems."

The third group of papers deals with projective techniques. There are two chapters on the Rorschach (one by Rorschach himself, drawn from the *Journal of Nervous and Mental Disease*, 1924) and two on the Thematic Apperception Test. They make interesting reading but they confirm the impression usually gained by psychologists who are not Projective experts, that these tests are highly subjective, have goals which change over the years and which are approached by devious, sometimes incompatible, routes.

Group 4 is categorized as "dimensional." This section includes A Blind Analysis of a Case of Multiple Personality using the Semantic Differential, by Osgood and Luria; and three characteristic chapters by R. B. Cattell, Eysenck and Rachman (23 references of which more than half are Eysenckian), and Hathaway and McKinley—the latter chapter being on the Minnesota Multiphasic Personality Inventory. The fifth, and last section comprises two chapters, one by A. L. Edwards, entitled Social Desirability and Personality Test Construction, the other by P. E. Vernon on the Concept of Validity in Personality Study. This is an abridgement of a chapter from Professor Vernon's recent book on *Personality Assessment*. It is an honest, gloomy piece, abounding in remarks like the following: "Neither of these claims has yet been confirmed by independent

investigations. But clearly such an approach, through patterns of scores, could have important bearing in counselling."

The general picture emerging from the book is that what psychologists have gained in methodological and experimental expertise, they have lost in psychological understanding and sense of humour. If the next few decades could recover the latter whilst retaining the former, the outlook would indeed by bright. A. W. HEIM.

The Psychology of Learning. By R. Borger and A. E. M. Seaborne. Harmondsworth, Middx.: Penguin Books (Pelican Original). 1966. Pp. 243. 5s.

The range of topics discussed in this introductory work is rather a large one for its size, for the authors choose to describe learning as "any more or less permanent change of behaviour which is the result of experience." Thus about one third of the book is devoted to animal learning and behaviour theory, while the remaining Chapters concern Language and concept formation; Memory; Perceptual learning; Skill; Abnormal behaviour; and Education. The resulting mixture is occasionally a little short of facts, but the clear and entertaining discussion of such a variety of topics will make the book a valuable introduction for the general reader. A well chosen bibliography gives a useful guide to more specialized works.

The scheme of the book is ambitious: to set out the important principles of learning discovered from experiments on animals, and to consider the ways in which these principles can be applied to various human problems, notably Education. In fact, however, the various sections of the book hang together rather badly, and the authors are not mistaken when they suggest in the Introduction that the topics in the middle Chapters (on Skill, Retention and so on) are poorly integrated with the discussion of animal learning. This is also sadly true of the section on Education, and although Borger and Seaborne argue vigorously for the application of the principles of learning theory to teaching, they give little evidence to convince the reader that such an application is useful at the moment.

What, after all, are the fundamental principles of learning supposed to be? The book's discussion of animal experiments does not suggest that there is any generally received opinion on this matter, and it would be badly misleading if it did. It is true that certain practical aspects of animal training are fairly well known to psychologists (as they are to circus trainers) but the principles involved do not take one very far in Education. Borger and Seaborne stress two such principles: the necessity for ensuring that the behaviour to be selected actually has some tendency to occur; and the need to reward it quickly when it does. These methods may be important, but it is hard to see their relevance to the main conclusion reached in the section on "The purpose of education," where it is argued that the communication of facts should be made secondary to the cultivation in pupils of the ability to learn. This old idea owes little to the work on animal learning, and the authors might have made it much clearer than they in fact have that their speculations on this matter have absolutely nothing to do with the rest of the book.

It is not surprising that there is such a gap in the argument, for several of the problems of education discussed in this book are just not scientific ones. The promise of a "significant breakthrough" lacks meaning because we are not told what it is that we are trying to break. The increased quantity of educational research advocated in the book will have no obvious function until we know what it is attempting to find out. Borger and Seaborne are aware that the prospect of large scale experiments being carried out on young minds will disturb some of their readers, but their claim that such experiments are justified by the necessity for progress seems to assume that we all agree on some final aim. It should be obvious that arguments on this matter will not be settled by any amount of work on the fundamental principles of learning. M. J. MORGAN.

Eliminating the Unconscious: A Behaviourist View of Psycho-analysis. By T. R. Miles. Oxford and London: Pergamon. 1966. Pp. xviii + 171. 17s. 6d.

Professor Miles has given us an interesting, challenging and courageous book. Its central thesis is that the traditional ways of presenting psycho-analysis embody the assumption of dualism between matter (or body) and mind. But "a dualist terminology is in no way an essential part of psycho-analysis"; and "sentences which purport to be about 'the mind,' 'the unconscious mind,' etc. can be translated without loss of meaning into sentences about behaviour." Behaviour includes the so-called "phenomenological sentences," such as "This looks green to me," "I have a toothache," "I feel very anxious

about her." For, on Miles's view, these sentences are signs that I have made certain discriminations.

In the latter half of the book Miles carries out this programme of logical analysis on the central concepts of traditional psycho-analysis. Its essence can be found in the chapter on the Unconscious. Suppose we argue, in a Freud-like way, that the expression "The Unconscious" refers to a real entity that we have to postulate as a hidden cause to explain the things which we do in fact observe—such as obsessive acts, panic fears, and the like. The sufficient objection to this argument is, "not that intangible causal agencies are necessarily an absurdity," but that such an explanation here is quite vacuous. For "the only evidence which can be adduced for the existence of 'the Unconscious' is the behavioural manifestations which require to be explained!" On the other hand, suppose we try to retain the concept by arguing that it is a façon de parler. We can then say *either* that the Unconscious is a "link concept"; *or* that it offers us a pictorial model. As a link concept, it would serve to connect, or link, the present symptomatic behaviour of the patient with other, especially earlier, events in his life. But "the fatal objection" to this view is that, for the purpose of linking, a substantive unconscious is "entirely unnecessary"; and it makes much better sense to speak here instead of "unconscious purposes" at work. Viewed as a pictorial model, the objection is that the model is usually supposed to be a picture of "the mind," which contains the Unconscious as a literal part. But this means that the model is logically tied to traditional dualism, and we are faced by questions like: "How can such an entity as 'the Unconscious' ever be known?" The fact is that the noun expression "the Unconscious" can be replaced without loss by the adjective "unconscious" or by the adverb "unconsciously," as they function in expressions such as "unconscious purposes," "wishes unconsciously," and so forth.

But do not expressions such as these commit us to accept the so-called "unconscious mental processes"? Miles argues that they do not; and he supports his argument by offering a Ryle-like and dispositional analysis of these expressions and related concepts. For example, suppose we say that a patient has unconscious knowledge, or knows unconsciously, that she is maltreating her child. Then what we are saying is that she knows this in the sense that she was acting as though she knows it—she knows it in the full sense, except that she is in no position to make a conscious avowal of it. On the constructive side, Miles maintains that what Freud and the analysts have done is to offer us a new family of concepts (e.g. unconscious knowledge, unconscious feelings), and a new orientation. These are important for the understanding "not merely of the two-person-therapy situation, but of many other social situations as well." Psycho-analysis serves as a "challenge to the psychologist to take the human relations factor seriously."

How sound is the case Miles presents? I think there are two basic objections to it. Let us suppose it is textually correct to say that Freud used the expressions under consideration dualistically. Then it is *not* correct to claim or assume (as Miles does) that the *only* way of eliminating the dualism is to eliminate the expressions, with their corresponding concepts. It is possible to remove the dualism by introducing and using the expressions in a way that is logically free from any dualistic involvements; but this is not the place to explain how this can be done. Moreover, it could be argued that, in speaking as he did, Freud was doing something very important. He was offering us—in a pioneering and groping way—an uninterpreted model of mental functioning; and this is very like what psychologists do today as a matter of course.

But Miles could reply that any explanation offered by Freud's model is vacuous. For the only evidence which can be adduced for the existence of the Unconscious are the symptoms which require to be explained. In order to link early history with present symptoms, the model of a substantive unconscious is entirely unnecessary. But now we meet the second basic objection to Miles's case. It is just false, or very misleading, to claim that the explanation is vacuous, and the model entirely unnecessary. For to explain the patient's symptoms as being due to, for example, a displaced manifestation of energy attached to unconscious mental elements, and so forth, is to direct our attention to the similarity between the patient's functioning and the workings of a system of elements built on the Freudian model. Far from this being vacuous, it may turn out to be of enormous scientific importance. Of course, it may be quite true that the explanation is vacuous *in the narrow sense* that it is not possible in principle to derive testable consequences from the model, which we cannot equally well derive without it. But is it not very misleading to overstress this fact? For it is quite sense to suppose that someone may come along shortly with a more precise version of the model, which *is* testable in principle and in practice. Suppose, furthermore, that this more precise model is actually

confirmed in a big way, and helps to revolutionize our whole way of regarding the detailed mechanics of human functioning. In retrospect we should then look pretty idiotic maintaining in the 1960s that the Freudian model is untestable, *and therefore* vacuous. So even if the model as it stands is untestable in principle at present, this feature may be relatively unimportant; and it is not sufficient to justify us now in rejecting the model as vacuous and entirely unnecessary.

It could be argued, therefore, that a much better defence can be offered of traditional psycho-analysis than Miles has allowed. Does this mean that Miles is mistaken in his whole programme to eliminate dualism from psycho-analysis, and to translate psycho-analytic sentences into sentences about behaviour? No, not at all. In so far as psycho-analytic discourse contains dualistic assumptions, Miles is doing something very important indeed in trying to eliminate them. Miles's programme of behaviourist translation forms an excellent exercise in therapeutic clarification for the psycho-analytic practitioner and for the psycho-analytically oriented psychotherapist. The book contains useful and important challenges to the psycho-analyst. It is to be hoped that it will not only be widely read in psycho-analytic and therapeutic circles, but that analysts will also make use of it in their own teaching programmes in their training institutes. B. A. FARRELL.

Manuel Pratique de Psychologie Expérimentale. By Paul Fraisse. Paris: Presses Universitaires de France. 2nd edition, 1963. Pp. 392. 20F.

In preparing a laboratory manual of experiments in psychology one must resolve two somewhat conflicting requirements. The first is to provide a series of concise experimental schedules with fairly detailed accounts of procedure and the second is to present enough of the theoretical background to make the experiments intelligible and interesting to students. A manual which concentrates too much on the former may reduce the experimental enquiry to a rather dull routine. On the other hand, an adequate treatment of theoretical principles may well inflate the size of the work to impossible proportions. Recent attempts to solve this problem have produced such widely differing approaches as that of the late Professor Humphrey in *Psychology through Experiment* and that of Mr. Babington Smith in *Laboratory Experience in Psychology*.

Professor Fraisse's manual follows a more traditional style, exemplified in England by the classic *Text Book of Experimental Psychology*, Part II, by Myers and Bartlett. The experiments have been chosen so that the balance between theoretical background and experimental procedure is well maintained over a wide range of topics in human experimental psychology, including some experiments on small groups. No work on animals is included.

The first edition of the Manuel appeared in 1956 and contained an introduction to experimental method followed by schedules for 60 experiments, and three appendices, one on the measurement of thresholds, a second on techniques of pen recording and a third giving a list of firms manufacturing the equipment used. All the experiments were relatively simple in conception, most could be carried out in 2 or 3 hr. on a single subject and the apparatus requirements were fairly modest. At the time of publication it was one of the most useful sources of practical class experiments available in this form.

In the new edition all the experiments of the first edition are retained with some minor improvements in exposition. Eight new experiments have been added together with a useful additional appendix on experimental design.

To have included in the new edition experiments on more recent developments, such as applications of information theory and signal detection theory, would have required a fairly lengthy theoretical treatment before the point of such experiments could be understood. It is perhaps for this reason that Professor Fraisse has chosen not to include such topics in the new edition (apart from a very brief allusion to information theory in an experiment on choice reaction times). In keeping with this policy the appendix on sensory thresholds remains unchanged and is confined to traditional methods for determining absolute and differential thresholds.

Although therefore one will not find experiments related to a number of developments which have taken place since the publication of the first edition, the manual retains its well balanced coverage of experiments of perennial interest to experimental psychologists and by publishing a second edition Professor Fraisse has kept a very useful set of workable experiments easily available in print. R. DAVIS.

Abstraction and Concept Formation. By Anatol Pikas. London: Oxford University Press (Harvard University Press). 1966. Pp. xiii + 303. 56s.

This interesting book, in which Pikas introduces his own recoding hypothesis, is more

remarkable for its breadth than its depth. Its merit is that it reviews in a critical way an immense amount of diverse research—the philosophical roots of the problem, the contribution of the Wurzburg and Leipzig schools, the clinical findings of Goldstein and Cameron, the symbol formation of Werner and Kaplan, the great classical experiments of Fisher, Hull, Heidbreder and Smoke, and finally mediation theories of learning such as those of Osgood and the Kendlers. The book also includes an excellent conceptual analysis of the usage of the terms "abstract" and "concrete." The account of the historical background is inevitably over-simplified as the author points out, but it contains some details of the greatest interest. "Ach also took his intelligent subjects to a department store, where he introduced the artificial words into the conversation, as denoting conceptual attributes in other objects. He was able to record that the Ss caught on immediately."

The style (or translation) is witty and engaging at first, but it deteriorates sadly as more complex theoretical issues are encountered. The writing begins to display frenzied, obsessional characteristics which would be censored in a Ph.D. thesis. "As regards the structure of this element (or perhaps we should say 'trait') in the primary code in the sphere of hypothetical constructions, its hypothetical existence is dependent precisely on the fact that a common response is given to dissimilar stimuli." There are also some startling naivities, e.g. the assertion that if the influence of Aristotelian logic can be traced in everyday life, then such a logic must have "always existed even before Aristotle and independently of him." There are statements which I fail to understand, e.g. that the distinction between inductive and deductive reasoning is synonymous with the distinction between conjunctive and disjunctive concepts. And there are one or two bad translations of which the worst is "memory track" for "memory trace."

Pikas's own hypothesis is based on Galton's "composite photograph" theory of memory. Instead of connections between codes a "melting together" or "summation" of codes is assumed to occur in concept formation. The author points out that he has not found any direct refutation of this theory—hardly surprising because it is difficult to know what could count against it. However, the following corroborative evidence is assembled which does suggest the "organic creation of recoding" in time. (1) Bourne's observation that a concept, once established, is resistant to misinformative feedback, and that there may be an optimal post-feedback interval. (2) An ingenious experiment of Rommetveit's which appears to demonstrate different qualitative stages in concept formation. (3) The Gestalt law of good figure and the phenomena of "normalization" demonstrated by J. J. Gibson. But, as the author admits, some of these phenomena may turn out to be consistent with some mediation hypotheses after all. This must be rather depressing for everyone. It seems to me that more crucial data may come from neurophysiological research—one thinks of Hebb at once.

It is interesting to note that Pikas is evidently not an experimentalist: no empirical observations of his own are quoted. The approach is typically inductivist. Instead of formulating a hypothesis and then doing experiments to test it, the author develops a plausible, global theory and then searches the literature for confirming evidence. However, it is always easy to be critical, and perhaps one should not be too captious when a man attempts a difficult task. This book, in spite of its faults, contains many good things, and it should be read by every research worker concerned with concept formation. It is likely to stimulate someone to do experiments of his own, to read the original source material, or at least to provoke self-critical discussion. My own impression is that Pikas has tried to do too much, and that this is the reason for the book's uneven quality. But I should be delighted to be proved wrong.

P. C. WASON.

Tactics of Scientific Research. By Murray Sidman. London and New York: Basic Books. 1966. Pp. x + 428. £1 7s. 6d.

This is a paperback re-issue of a book first published in 1960. Belying its title, it is not a primer of statistical method. On the contrary it is a leisurely, lengthy essay on the rationale underlying the approach to the study of behaviour identified with the name of Skinner, and as such precisely *not* a primer of statistical methods. Although somewhat discursive, it succeeds well in presenting the justifications for a Skinnerian approach; indeed, since Sidman does not devote time and effort to insisting on response-rate as the only rational measure of behaviour, he provides a welcome change from the writings of Skinner himself on this general topic. Much of what Sidman has to say (e.g. on variability and replicability) is interesting, important and well worth the attention of the most devout opponent of Skinner's system. It is only when he elevates his tactics

into a binding strategy that doubts enter. For example, much disapproval is voiced against the use of group designs: the effects of any variable are to be studied by repeatedly changing the value of the variable in question in experiments with single subjects. Such a simple strategy, of course, assumes that the behavioural changes so induced are reversible—frequently an implausible assumption. Instead of accepting the use of a group design in such cases, Sidman wastes much ingenuity in attempting (with little success) to show how to solve the problem with a single-subject design. Single-subject designs are often better than group designs; but to insist that one never uses the latter is severely to restrict the range of problems tackled. As Sidman says "There are no rules of experimental design. . . . In our search for new information we must be prepared at any point to alter our conception of what is desirable in experimental design" (p. 214). If this is true, it is not compatible with adherence to some of the tenets of Skinnerian dogma.
N. J. MACKINTOSH.

Psychodynamics and Hypnosis: New Contributions to the Practice and Theory of Hypnotherapy. Compiled and Edited by Milton V. Kline. Springfield, Illinois: Thomas. 1966. Pp. xi + 194. $8.75.

This indifferent volume contains seven papers concerned with various aspects of hypnosis and its uses in psychotherapy. Apart from a paper by the Editor, its interest to readers of this *Journal* is negligible.
O. L. ZANGWILL.

Psychology: The Science of Mental Life. By G. A. Miller. London: Penguin Books. 1966. Pp. 415. 7s. 6d.

This is a re-issue of a book first published in the U.S.A. in 1962 and subsequently in this country by Hutchinson. It evolved out of a series of introductory lectures given at Harvard, and so, in the new cheap format, must be considered primarily as an introductory text. As such it can be confidently recommended, for it is among the small number of books that give a balanced and above all an historical perspective on current psychological thought. Miller uses biographical essays on six pioneers (Wundt, James, Galton, Pavlov, Freud and Binet) to introduce the modern ideas which they initiated. For example, the essay on Wundt is followed by chapters on various aspects of consciousness, that on Pavlov by a discussion of learning theories and animal behaviour, and that on Freud by chapters on motivations and drives. The value of this approach lies not in any comprehensive study of the entire field of modern research, but rather in a survey of major lines of thought. In addition, the growth of each man's ideas is related to his intellectual milieu. It is here that the strength of the book lies. With long, confident steps Miller leads us through the jungle of nineteenth-century "isms" out of which the new psychology grew. Particularly successful are his chapters on Wundt and James, whose battles he describes in a style more literal than metaphoric; throughout the book his thought is clear and his writing vigorous.

Penguin addicts will find here a balanced view of the subject, which they have needed particularly since the publication of the Penguin Modern Psychology Series. This is both an excellent, cheap introductory text and a complementary volume to the large number of more specialized paperbacks now available. However, this must not be regarded solely as beginners' reading, for, as Brian Foss says in his editorial forward, "Miller is able to describe psychology so that both beginners and jaded old hands have their appetites whetted."
C. J. DARWIN.

The Memory System of the Brain. By J. Z. Young. London: Oxford University Press. 1967. Pp. vii + 128. 28s.

This book is based on a series of lectures delivered in 1964 at the University of California. It provides a summary of some of the experimental findings and the more important concepts described in Young's *A Model of the Brain* Oxford: Clarendon Press. 1964.

Professor Young and his collaborators have spent some twenty years investigating those central mechanisms that govern behaviour in the octopus. The gross anatomy of the central nervous system in cephalopods is quite different from that of vertebrates, which is scarcely surprising since philogenetic development of these two organs has proceeded along independent lines for roughly 500 million years. Nevertheless, similar conditioned reflexes can be established in both dog and octopus, and the latter often learns faster. Young has always hoped that exploration of two systems that were so different in gross structure might reveal those factors at the cellular level responsible for all learning.

He says "The octopus system, by its relative simplicity, reveals certain features that we may look for in higher nervous systems."

He points out that, in its simplest form, learning consists in the establishment of a choice between two alternative motor responses to a particular form of sensory stimulation. "In the octopus we have two such situations: the attack or retreat when some object moves in the visual field, and the taking or rejecting of an object by the arms." These activities can readily be conditioned to visual and tactile cues. Most of the recent work described in this book, has been concerned with disturbances in memory and learning ability, inflicted by removal of various lobes from the nervous system. For instance, it has been shown that there are centres, essential to visual and tactile memory, which '. . . can be removed separately without influencing each other."

Professor Young has developed a working hypothesis which postulates that learning is determined by a functional unit, consisting of a group of several cells that he calls a *mnemon* (see also Young, *Proc. Roy. Soc. B*, **163**, 285). These are the decision-making units of the nervous system, governed by sensory feedback informing them of the desirable or undesirable consequences of recent activity. Young assumes that the output of such units is arranged to produce inhibition of an "unwanted pathway." Incoming sensory information must be filtered through many such mnemons before reaching the animals' motor outflow. "It is assumed that the inappropriate channels from the cells stimulated are fully closed by the signals of results arriving at the learning system, and that they do not reopen . . . . The familiar slow development of the evidence of learning consists in the accumulation of enough trained cells to ensure a consistent response."

This is a stimulating book, but not nearly so good as its parent, *A Model of the Brain*. I believe that it would be hard for a reader, who had not seen other works by Young, to disentangle fact from hypothesis. For example, while the mnemon was doubtless conceived as a result of the work in Naples, it does not appear to be established or specifically supported by the experimental results. Nevertheless, the description of this concept is dispersed among accounts of experimental observations in such a way that the unwary reader might believe it to be more than a useful working hypothesis. This unintentional confusion of fact with fancy is doubtless the consequence of trying to say too much in too little time. But the result is a book which does not do justice to Young's important contributions in a fascinating field of study.

B. Delisle Burns.

PUBLICATIONS RECEIVED

BOOKS

Review in a later issue is not precluded by notice here

The Uptake and Storage of Noradrenaline in Sympathetic Nerves. By Leslie L. Iversen. London: Cambridge University Press. 1967. Pp. xiv + 253. 57s. 6d. $11.00.

Functional Teaching of the Mentally Retarded. By Max G. Frankel, F. William Happ and Maurice P. Smith. Springfield, Illinois: Thomas. 1966. Pp. xvii + 241. $9.75.

The Psychiatric Halfway House: A Case Study. By Naomi D. Rothwell and Joan M. Doniger. Springfield, Illinois: Thomas. 1966. Pp. xv + 265. $8.75.

A Workbook in the Rorschach Technique Emphasizing the Beck and Klopfer Systems. By John E. Exner. Springfield, Illinois: Thomas. 1967. Pp. viii + 113. $10.50.

Measurement and Prediction. By S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star and J. A. Clausen. London and New York: Wiley. Science Editions 1966. Pp. x + 756. 23s. ($2.95).

Porteus Maze Test: Fifty Years Application. By Stanley D. Porteus. London: Harrap. 1966. Pp. 320. 55s.

Explorations in Typing Psychotics. Edited by Maurice Lorr. London and Oxford: Pergamon. 1966. Pp. viii + 241. 60s.

Constancy and IQ Change. By Alice E. Moriarty. Springfield, Illinois: Thomas. 1966. Pp. vii + 233. $8.50.

Clinical Psychopharmacology of Anxiety. By J. R. Wittenborn. Springfield, Illinois: Thomas. 1966. Pp. x + 228. $6.75.

Children: Development and Relationships. By Mollie S. Smart and Russell C. Smart. London: Collier-Macmillan. 1967. Pp. ix + 582. 63s.

REVUE CANADIENNE CANADIAN JOURNAL DE PSYCHOLOGIE OF PSYCHOLOGY

VOLUME 20, NO. 4 DECEMBER/DÉCEMBRE 1966

Published for the Canadian Psychological Association by
PUBLIÉ POUR LA SOCIÉTÉ CANADIENNE DE PSYCHOLOGIE PAR
UNIVERSITY OF TORONTO PRESS

THE CANADIAN JOURNAL OF PSYCHOLOGY is published bi-monthly. Annual subscriptions, $10.00, single number $2.00.

Contributions. Original manuscripts and correspondence on editorial matters should be sent to:

The Editor, Canadian Journal of Psychology, McMaster University, Hamilton, Ontario.

Subscriptions to Journal. Orders and correspondence regarding subscriptions, changes of address, and purchase of back numbers shall be sent to:

Canadian Psychological Association,
The Secretary-Treasurer,
Suite 210, 225 Lisgar Street,
Ottawa 4, Ontario.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIX 1967 Part III

SOME OBSERVATIONS ON GREGORY'S THEORY OF PERCEPTUAL ILLUSIONS

BY

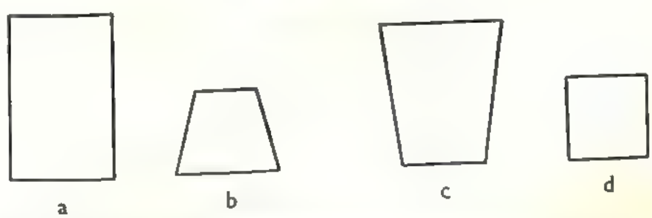
MARIO ZANFORLIN*

From the Instituto di Psicologia Sperimentale, University of Padua

Gregory (1963, 1965, 1966) has explained the optical illusions of the Müller–Lyer type as a misapplication of a perceptual mechanism which he calls "primary constancy scaling." This mechanism is brought into action by the depth features of the drawing, despite the fact that the texture of the paper makes the observer see the figure as flat. It is from the conflicting information about the figure, namely, the depth features (three-dimensionality) and the texture (flat), that the well-known phenomenon of these illusions arises.

As Gregory's replies to some criticisms (Brown and Houssiad, 1964; Humphrey and Morgan, 1965; Day, 1965; Hamilton, 1966; Wallace, 1966) do not appear to satisfy all doubts about his theory, it is opportune to consider in detail certain basic points of the theory which have not been raised by previous critics. It seems obvious that if primary constancy scaling is set by the depth features of flat figures (Gregory, 1963), then in order to predict in which figures constancy scaling will enter into action and produce illusion, we need an exact definition of depth features. But no formal or precise definition is given, nor does Gregory determine unequivocably the circumstances in which they occur. What depth features are is not at all self-evident. Gregory (1963) speaks of a drawing as having depth features when it is a perspective representation or a projection of a three-dimensional object. This does not help us very much as any drawing whatsoever may be regarded as the projection of at least one other figure in "depth." If, for example, one supposes that Figure 1*b* has depth

FIGURE 1



features because the drawing is the perspective of rectangle 1*a*, then one must also suppose that the square too of Figure 1*d* has depth features because it is the perspective of Figure 1*c*, although Gregory (1965), it seems, does not consider a square as having depth features. Nor is it possible to solve the problem by making

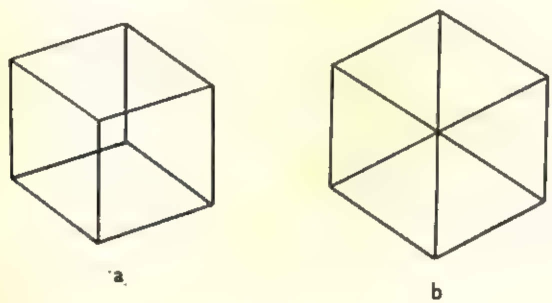
* Now at Department of Zoology, University of Edinburgh.

a distinction between "typical" and "non-typical" projections of objects, since no criteria are given on which to base the distinction. We cannot therefore consider perspective as a definition of depth features.

Gregory's (1965) methods of establishing depth features by presenting self-luminous figures in the dark to see if the subjects perceive apparent distance, and thence deducing whether the figure does or does not have depth features, does not indicate what depth features are, and it allows dubious predictions.

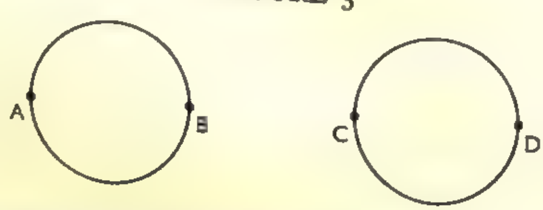
In Figure 2, for example, there are two projections of a cube (Kopfermann, 1930), both have oblique lines, and both have an equal number of edges; but in Figure *2a* the drawing is seen as three-dimensional, whereas in Figure *2b* it appears flat, and it makes no difference if the figures are drawn on paper or made self-luminous. In which of these two figures are depth features present and what are they?

FIGURE 2



In an attempt to make a prediction, let us consider the two circles of Figure 3. When they are made self-luminous they appear as two equal circles at the same distance in depth (Zanforlin, unpublished experiments; modifications of illusions to be found in Titchener, 1905). Is the conclusion that there are no depth features? It is true that there is no apparent distance and so it would seem that there are no depth features, and hence no illusions in this figure; but it is also true that this figure does give rise to a remarkable illusion. (The point B does not appear to be at an equal distance from point A and point C.) And so one could maintain that depth features are present as we still do not know what these are.

FIGURE 3

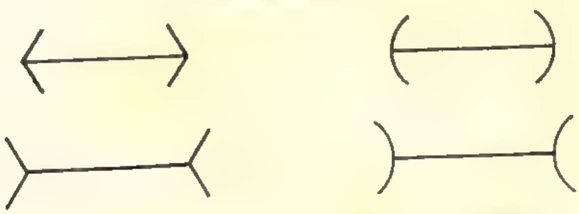


In the case of the Müller–Lyer illusion, Figure *4a*, Gregory (1963, 1965, 1966) very clearly points out that with self-luminous figures the "arrows" make the two lines appear one nearer than the other. If we remove the arrows, the two equal lines will appear at the same distance. So the arrows are supposed to be the depth features. If we now draw two simple lines one shorter than the other as in Figure 5, the shorter of the two will appear further away and the bigger one nearer (Zanforlin, unpublished experiments). What is it now that causes the apparent distance when no arrows-depth-features are present? One possible explanation is that Figure 5 is the

representation in perspective of two equal lines at different distances. But as far as perspective is concerned, Figure 5 is also the projection of two lines of different length at the same distance. So we would expect at least 50 per cent. of the subjects to perceive this second possibility. But this is never the case with self-luminous figures, as all the subjects do, in fact, perceive the two lines as more or less equal and at different distances. So having excluded perspective and "arrows" as explanations for the apparent distance in self-luminous figures we do not know what brings constancy scaling into action in the case of, for example, the illusion of parallels in Figure 6 (Titchener, 1905).

From the above considerations, it is clear that the theory does not allow exact predictions as to which figures will give rise to illusions. But in one particular case Gregory gives us an interesting experimental result that makes an exact prediction possible.

FIGURE 4



a b

FIGURE 5

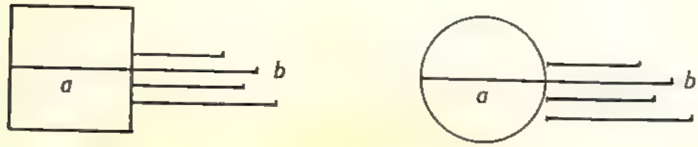
FIGURE 6

In the case of Figure 7 in which the "arrows" are at 90° (vertical) to the horizontal segment, according to the results of Gregory's (1965) experiments there are no depth features, because with self-luminous figures there is no apparent distance. So we can predict that no illusions are possible with vertical "arrows." This prediction, however, can be contradicted by the Oppel illusions (Boring, 1941) where some more vertical "arrows" are added (see Fig. 8). Moreover, if two horizontal lines are added to Figure 7 (are we adding depth features?) in such a way as to make a square, Figure 9*a*, we obtain another illusion (Titchener, 1905; Zanforlin, unpublished). The line *a* in the middle of the square appears to be shorter than the external segment *b* despite the fact that they are both equal, the same illusion is obtained in the case of the circle, Figure 9*b*, where the diameter appears to be shorter than the external segment *b*. There can be little doubt that these illusions are not explicable on the basis of primary constancy scaling since they have no depth features.

FIGURE 7

FIGURE 8

FIGURE 9



a b

But Gregory's fundamental hypothesis it seems, does not depend on constancy scaling nor on depth features, but on the assumption that these illusions are perceived only when we find ourselves presented with drawings (Gregory, 1963, 1966). It follows from this that when we look at real objects in the world, we should never have illusions of this type. The following experiments show this assumption to be untenable. Figure 10 shows two wooden mats and two sticks. One of the sticks is placed so as to form the diameter of one mat and the other is placed between the two mats. The sticks are equal in length, but every subject tested declared without hesitation that the stick in between two mats appeared to be longer. The phenomenon does not change if instead of "flat" wooden mats, two equal balls are used. If a subject is asked to make a pile of coins equal in height to the diameter of the coin (Figure 11) the pile, when measured, will always be about 30 per cent. too low (Metzger, personal communication).

There is no doubt that our eyes do "deceive" us, not only when we look at drawings of the Müller–Lyer type, but even when we look at real objects, for example, the books on our desk.

There is one further observation to be made with regard to "circular illusions," Gregory (1963, 1966) believes that a "cultural" influence (namely the fact that we live and are reared in houses with corners) plays a part in the perception of these illusions. As one piece of evidence, he quotes the research of Segal, Campbell and Herskovits (1963) on primitive peoples who are thought to have a "circular culture" because they live in round huts with no corners. These people, according to Segal *et al.* are not subject to illusions of the Müller–Lyer type with figures that have corners (e.g. Fig. 4*a*). The authors ought perhaps to have presented to the primitive people not only figures with corners but also figures with curves, as in Figure 4*b*, before drawing any conclusions about cultural influence. For if we assume that the type of culture determines the types of illusions, then people that have a "circular culture" must have illusions in figures with curves like Figure 4*b* as we have illusions in figures with corners. On the basis of this assumption it is also very difficult to explain why "we," despite living among corners, are subject to illusions produced by "circular" figures.

But leaving primitive and civilized people aside, one should remember that it has been proved (Warden and Barr, 1929; Winslow, 1933) that even birds are subject to illusions of the Müller–Lyer type and it is difficult to assume that their illusions derive from the fact that their nests have "corners."

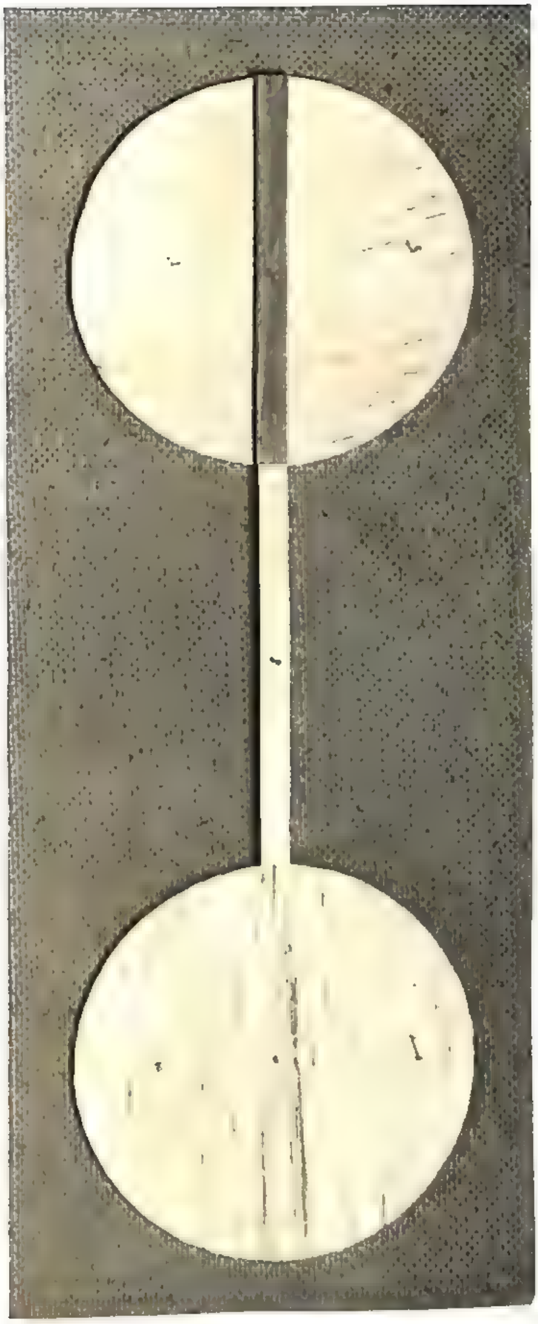
I would like to thank Dr. A. Manning, Professor F. Metelli, Professor W. Metzger and Professor R. C. Oldfield for having read the manuscript and corrected my English.

References

Boring, G. E. (1941). *Sensation and Perception in the History of Experimental Psychology.* New York: Appleton-Century-Crofts.

Brown, L. B., and Houssaid, S. L. (1964). Illusory perception as a constancy phenomenon. *Nature*, **204**, 302–3.

FIGURE 10



The Müller-Lyer illusion produced with solid objects: two wooden mats.

To face p. 196

FIGURE 11

Illusion with solid objects: the coins' illusion.

DAY, R. H. (1965). Inappropriate constancy explanation of spatial distortions. *Nature*, **207**, 891–3.

GREGORY, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, **199**, 678–9.

GREGORY, R. L. (1965). Seeing in depth. *Nature*, **207**, 16–19.

GREGORY, R. L. (1966). *Eye and Brain.* London: Weidenfeld & Nicolson.

HAMILTON, V. (1966). Susceptibility to the Müller–Lyer illusion and its relationship to differences in size constancy. *Quart. J. exp. Psychol.*, **18**, 63–72.

HUMPHREY, N. K., and MORGAN, M. J. (1965). Constancy and the geometric illusions. *Nature*, **206**, 744–5.

KOPFERMANN, H. (1930). Psychologische Untersuchungen über die Wirkung zweidimensionaler Darstellungen körperlicher Gebilde. *Psychol. Forsch.*, **13**, 293–364.

SEGAL, M. H., CAMPBELL, D. T., and HERSKOVITS, M. J. (1963). Cultural differences in the perception of geometric illusions. *Science*, **139**, 769–71.

TITCHENER, E. B. (1905). *Experimental Psychology.* London: Macmillan.

WALLACE, G. K. (1966). Optical illusions. *Nature*, **209**, 327–8.

WARDEN, C. J., and BARR, J. (1929). The Müller–Lyer illusion in the Ring Dove *Turtur risorius*. *J. comp. Psychol.*, **9**, 275–92.

WINSLOW, C. N. (1933). Visual illusions in the chick. *Arch. Physiol.*, **153**, 1–83.

Manuscript received 3rd April, 1967.

CONTRAST AND CONFLUXION AS COMPONENTS IN GEOMETRIC ILLUSIONS

BY

VEIJO VIRSU

From the Institute of Psychology, University of Helsinki, Finland

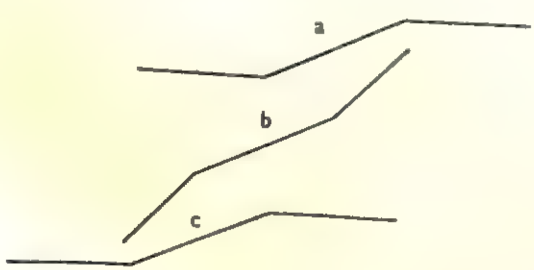
Two experiments on illusions were planned to test the predictive power of explanations based on size contrast and confluxion. The predictions turned out to be correct. A modification of the Müller–Lyer figure and an illusion of divided distance were used as stimulus figures. In the latter the dividing distances were underestimated. It proved to be methodologically necessary to measure contrast effects as differences between results of two experimental procedures, because small intervals as such were overestimated. It was not possible to explain all the results by means of the constancy theory in the form suggested by Gregory. Neither contrast nor constancy alone is sufficient to explain the geometric illusions. The development of an adequate theory by combining these two explanations seems possible.

Introduction

During the last 15 years almost all the research connected with theories of geometric illusions has been related to perspective theory. Tausch (1954), Kristof (1961) and Gregory (1963, 1965) have developed this type of explanation in the form which is here called the constancy theory. Its fundamental assumption is that certain flat figures suggest depth. The mechanisms of size and shape constancy are activated, which leads to nonveridical flat figure perception. As far as size proportions are concerned, it may be predicted on the basis of the previous assumption that those parts of the figure are overestimated which, according to the depth cues, could be localized further in space (primary scaling), or alternatively, those parts of the figure which appear to be at a greater distance are overestimated (secondary scaling; Gregory, 1963).

The constancy theory leads to better predictions than the other explanations previously given. However, the constancy explanation of some illusions, as for example Lipps's illusion (Fig. 1), seems unnatural, and it is doubtful whether the constancy theory succeeds here at all. In Lipps's figure the parallel lines *a*, *b* and *c* do not look parallel.

FIGURE 1



The optical illusions also have haptic analogies (see e.g. Cook, 1909; Révész, 1934; Rudel and Teuber, 1963) and to explain these haptic analogies by reference to the constancy mechanisms proves difficult. One explanation which seems to avoid these

difficulties dates back to Helmholtz and Mach. It is based on contrast and confluxion, and we shall call it here the contrast explanation.

Generally speaking the contrast effect, or briefly contrast, in perception is a phenomenon in which differences between clearly different stimuli of the same modality are accentuated. Confluxion can be regarded as a phenomenon in which the perceived differences between similar stimuli are diminished. Contrast and confluxion are parts of the same continuum: confluxion is, in a way, negative contrast. We are concerned primarily with figural size proportions. The concept of confluxion implies, first, that those parts of the figure which extend close to other parts inside the figure, or which are parts in relatively small subfigures, are enlarged, and secondly, that those parts which are only slightly smaller than parts of the figure they are compared with are enlarged. The former is intrafigural confluxion and the latter interpartial confluxion (cf. Cymbalistyj, 1949). These types of confluxion are not mutually exclusive. The reverse is implied by the concept of contrast: those parts inside the figure which do not extend close to other parts, or which belong to relatively large parts of the figure, diminish (intrafigural contrast); the same happens to parts considerably smaller than the parts which they are compared with (interpartial contrast). If the discrepancy is extremely great, no size contrast occurs (Ebbinghaus, 1913, p. 65). We suppose that contrast and confluxion are symmetrical phenomena. This implies that if A is contrasted with B and A is smaller than B, A is underestimated and B overestimated. Similarly, if a confluxion of A and B takes place, and A is smaller than B, A is overestimated and B underestimated.

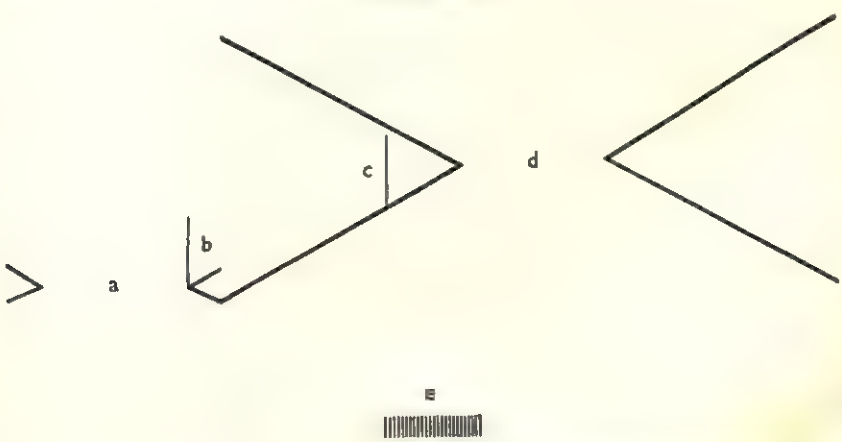
The two experiments described below were designed to test predictions based on the contrast explanation.

Method

Experiment 1

Figure 2 was presented to 24 undergraduates, half of whom were students of psychology.

FIGURE 2



In Figure 2 the distances *a* and *d* were 40 mm., and both the lines *b* and *c* were 20 mm. long. The shafts *a* and *d* have been omitted in order to avoid errors in measuring the distance between the angles (fins). Part *e* in Figure 2 represents a millimeter scale. The stimulus figure was drawn on a piece of thin white cardboard, 16 × 24 cm. in size. It was viewed from a distance of 50 cm. The subjects were asked to judge whether the distances *a* and *d* were of the same length or not. The distances were indicated to the subjects by pointing because there were no letters in the original stimulus figure. If the distances were judged different, the subject was asked to estimate the difference in

millimeters using scale *e*. He then judged the length of the lines *b* and *c* in a similar way. (Half the subjects judged the lines first.) When the subjects had judged the lines they were asked to think of the lines in the figure as projections of three dimensional figures, and they were instructed to give the order of apparent distance of *b* and *c* as well as of the imagined lines *a* and *d*.

Experiment 2

Procedure A. Twelve stimulus figures corresponding to Figure 3A were drawn on white cards, 16 × 24 cm. in size. The midpoint of the figure was located upwards to the right, 1 cm. off the midpoint of the card. In each of the 12 cards the total length *ab* of the stimulus figure was 100 mm. The height of the dividing bars was 14 mm. and their breadth 0·7 mm., each. The density and the distances between the midpoints of the bars varied as indicated in Table I. The distance *cd* between the bars expressed in millimeters equals the percentage the bar interval ("dividing distance") is of the total figure ("divided distance"). The letter F indicates a black rectangle 100 mm. in length and 14 mm. in height.

TABLE I

NUMBER OF DIVIDING BARS (NAME OF FIGURE) AND DISTANCE BETWEEN THE BARS (IN MM.) IN THE FIGURES USED IN EXPERIMENT 2

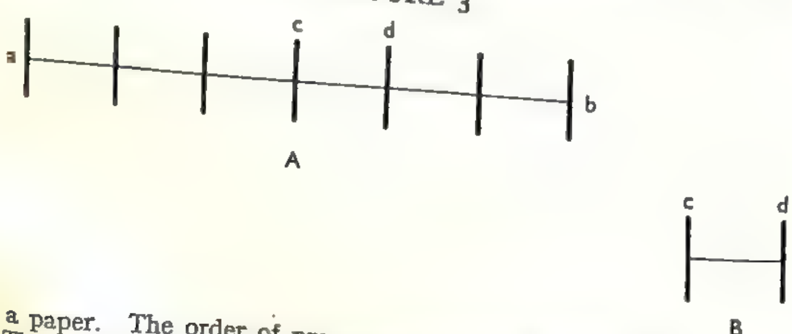
Number of dividing bars	0	1	3	5	7	9	11	13	17	24	49	F
Distance between bars	100·0	50·0	25·0	16·7	12·5	10·0	8·3	7·1	5·5	4·0	2·0	0

On the presentation of each stimulus figure the subject was asked to mark on a separate paper two points to indicate the apparent length of the thin line *ab*, and then mark between his markings the length of the line connecting *c* and *d*. The four markings for each stimulus figure were made on a separate response sheet, 14·8 × 21·0 mm. in size. In the case of the Figures O and F only two points were marked. The stimulus figure was 55 cm. from the subject. The response sheet was placed on the table on a grey background approximately 10 cm. to the right and closer to the subjects than the stimulus figure. The order of presentation was determined by two randomly chosen 12 × 12 Latin squares. To the subjects of the latter square, Figure O was presented together with Figure 3B when the distance *cd* was 10 mm. Similarly Figure F was presented together with a corresponding figure when the distance between *c* and *d* was 50 mm.

Twenty-four subjects were employed of whom 16 had studied psychology.

Procedure B. The stimuli corresponding to Figure 3B were presented to the subjects as in Part A of the experiment. The number of stimuli was 11. The distances *cd* were the same as in the previous experiment with the exception of Figure F which was omitted. The subjects expressed their judgements of the apparent length of *cd* by marking two

FIGURE 3



points on a paper. The order of presentation was determined by two 11 × 11 Latin squares. The subjects of the first square marked the points on blank paper. Those of the second square marked the distance *cd* between two points which were on the response sheet. The distance between these two points was the same as the average apparent length *ab* in each figure obtained using Procedure A.

Twenty-two subjects were tested; 10 of whom studied psychology. Five subjects participated in both parts of this experiment.

Predictions

Figure 2 is a modification of a figure originally introduced by Ebbinghaus (Ebbinghaus, 1913, p. 64). The Ebbinghaus figure was used by Kristof in one of his experiments (Kristof, 1961, p. 43). The result obtained by Kristof was that the distance *a* was judged longer than the distance *d*. The result was interpreted by Kristof according to the perspective theory: the fin figure *a* is localized further back because it is smaller than the fin figure *d*. The figure used in Kristof's experiment allows this interpretation; the explanation is, however, doubtful. Owing to the apparent depth in Figure 2 the distance *d* is perceived further back than *a*. The contrast explanation leads us to predict that the distance *d* will be perceived still smaller than *a*, because there is interpartial contrast between *d* and the big fins; and also *a* will be perceived as greater due to interpartial confluxion with the smaller fins. An underestimation of bar *b* as compared to *c* should occur, because the position of *b* is liable to intrafigural contrast, and *c* to intrafigural confluxion. On the basis of the constancy theory it may be predicted that the distance *d* will be overestimated as compared to *a*, and bar *c* in turn overestimated as compared to *b*, if secondary scaling operates in the supposed direction.

Figure 3A is a modification of Oppel's figure. It has been established that a divided line *ab* is overestimated. The constancy theory does not provide an explanation. This is seen clearly if the elements to be compared are two squares with stripes in opposite directions (Helmholtz's squares). According to the contrast explanation interpartial contrast between the total distance *ab* and the small intervals *cd* occurs. This implies that distance *ab* is perceptually enlarged and distance *cd* perceptually diminished. In one of the experiments by the present author it turned out that the Oppel's illusion was a parabolic function of the number of the dividing bars (the least square parabola was $E = -0{\cdot}052\ D^2 + 0{\cdot}980\ D + 1{\cdot}761$). The illusion reached its maximum with eight bars. In that experiment the breadth of the bars was 1/20 of the divided distance, the number of dividing bars varied from 0 to 19 and the number of subjects was 20. The method was similar to that reported in Virsu (1966). Also the bisected distance was overestimated. It is generally assumed, however, that underestimation occurs in this situation (see e.g. Oyama, 1960). An underestimation of the bisected distance cannot be understood on the basis of the constancy theory. The contrast explanation makes both results intelligible. According to Cymbalistyj (1949) the whole-perceiving attitude is favourable to confluxion. Depending on the instructions and experimental conditions either contrast or confluxion takes place in a figure divided with a single bar because confluxion may occur with the great dividing distance. Contrast occurs if the dividing distance is smaller than half of the divided distance. The contrast effect increases first but begins to decrease when the size difference between the dividing distance and the total distance becomes very great. On the basis of the constancy theory it is plausible to assume that the maximal constancy effect is reached with large dividing distances.

Taking symmetry into consideration it may be predicted that in Figure 3A the dividing distance will be underestimated. It may also be predicted that the underestimation of the dividing distance is a curvilinear function of the size of the dividing distance; the function is supposed to be first an increasing and then a decreasing function in the same way as the size of the illusion. If a prediction is to be made in this situation on the basis of the constancy theory, we should assume that the dividing distances are overestimated.

Results

Experiment 1

The subjects judged, on the average, the distance *a* to be 2·71 mm. or 6·77 per cent. longer than *d*. Line *c* was assumed on the average to be 1·71 mm. or 8·55 per cent. longer than line *b*. The standard deviations of the estimations were 2·31 and 1·06 mm., respectively. In both cases the difference between the mean error of estimation and zero is highly statistically significant. Six subjects judged *a* equal to *d* and 18 subjects judged *a* longer than *d*. Correspondingly, *c* was judged equal to *b* by four subjects and longer than *b* by the other 20.

The order of perceived distance in depth is indicated in Table II.

TABLE II

The order from the nearest to the most distant part of the figure	*Frequency*
a = b, c, d	12
a = b = c, d	1
a, b, c, d	3
b, a, c, d	1
b, c, a, d	1
d, c, a, b	1
c, d, b, a	1
d, c, a = b	4

The equality sign indicates a judgement of equal distance.

The result is in accordance with the prediction based on the contrast explanation. On the other hand, the predictions of the constancy theory for the distances *a* and *d* were wrong. Only four subjects gave responses placing *a* and *d* in non-extreme positions. On the basis of this result it is plausible to assume that secondary scaling can take place when the subject is judging this figure. The four subjects who perceived the order of distances contrary to expectation showed errors of estimation in the same direction as the rest of the subjects, except one subject who judged the bars *b* and *c* to be of equal length.

Experiment 2

The results obtained by Procedure A can be seen in Tables III and IV. A negative sign means underestimation. Table III shows the means and standard deviations of the estimation errors of the distance *ab* expressed in mm. and in percentages. Table IV shows the corresponding values for *cd*.

TABLE III

Errors in the Reproduction of the Divided Distance *ab* in Figure 3A

Figure	0	1	3	5	7	9	11	13	17	24	49	F
Mean mm. and mean per cent.	−3·42	−3·17	−0·79	−0·33	0·12	2·25	2·29	0·67	2·17	1·75	−1·62	−8·67
S.D.	6·20	8·10	7·04	8·76	8·75	7·75	7·18	7·71	7·60	10·39	8·74	8·18

TABLE IV

ERRORS IN THE REPRODUCTION OF THE DIVIDING DISTANCE *cd* IN FIGURE 3A

Figure	1	3	5	7	9	11	13	17	24	49
Mean mm.	−5·35	−0·66	0·34	1·42	1·27	1·51	1·05	0·87	0·65	0·47
Mean per cent.. ..	−10·7	−2·6	2·0	11·4	12·7	18·1	14·7	15·7	16·2	23·5
S.D. mm.	6·97	3·20	3·00	3·58	1·91	1·66	1·24	1·51	0·98	0·91
S.D. per cent.	13·9	12·8	18·0	28·6	19·1	19·9	17·4	27·2	24·5	45·4

TABLE V

ERRORS IN THE REPRODUCTION OF THE FIGURE 3B

Figure ..	0	1	3	5	7	9	11	13	17	24	49
Mean mm. ..	−1·28	0·00	2·68	2·83	2·93	1·90	1·77	1·79	1·25	1·14	0·68
Mean per cent. ..	−1·3	0·0	10·7	16·9	23·4	19·0	21·2	25·1	22·5	28·5	34·0
S.D. mm.	7·13	4·48	2·49	1·77	1·68	1·36	1·11	1·36	0·84	0·92	0·62
S.D. per cent. ..	7·1	9·0	10·0	10·6	13·4	13·6	13·3	19·0	15·1	22·9	30·9
Square I mean mm.	−4·48	0·06	3·28	2·76	3·09	1·87	1·46	1·63	1·36	1·15	0·61
Square II mean mm.	1·93	−0·06	2·08	2·90	2·77	1·94	2·08	1·95	1·14	1·14	0·75

The results obtained by Procedure B are given in Table V. The means of the Latin squares I and II show only small differences which can be regarded as chance effects. It is concluded that the marks the subjects made in part A of the experiment to indicate the distance *ab* have no effect on the estimations of the distance *cd*. The greatest square difference was obtained with the Figure O, where *cd* is 100 mm. The result of the square II may be explained by referring to the two-point drawings on the response sheets which in the majority of cases exceeded the distance of 100 mm. Possibly the subjects supposed that the distances of these points were same as the stimulus distance of 100 mm.

The main results of Experiment 2 are shown in Table VI and in Figure 4. The average contrast effect can be obtained by subtracting the mean errors of estimation in Procedure A from the mean errors obtained using Procedure B for the estimation of the same distance *cd*. Errors due to the experimental procedure ("unexplained errors") can be eliminated this way, and the existence of contrast is seen in positive values. Table V shows the degree of distortion in the stimulus figure when no other dividing bars are present. The results indicate a new type of illusion: the distance between two vertical lines which are near each other may be overestimated by as much as 30 per cent. of the physical distance. With large differences in size the contrast phenomenon may even change to the opposite, at least when subjects are instructed to estimate apparent size. The results in Table IV show the error in the distance *cd* when the predicted contrast effect and the error due to the experimental procedure are confounded. This type of mixed effect in investigating geometric illusions is often interpreted as the final effect. In this case the error measured differs from that supposed to be measured.

The results in Table VI are clearly in the predicted direction. In the second Latin square in part A the Figures 1 and 9 which correspond to Figure 3B showed a mean effect of contrast of 6·09 and 0·72 mm., respectively. This result corresponds to the result obtained by using the procedures A and B together. It may be noted that each of the five subjects who participated in both parts of the experiment had a higher total mean of errors in part B than in the corresponding task in part A. Only one of these five served as subject in the second square in part A. The correspondence between the experimental groups can be considered fairly good.

TABLE VI

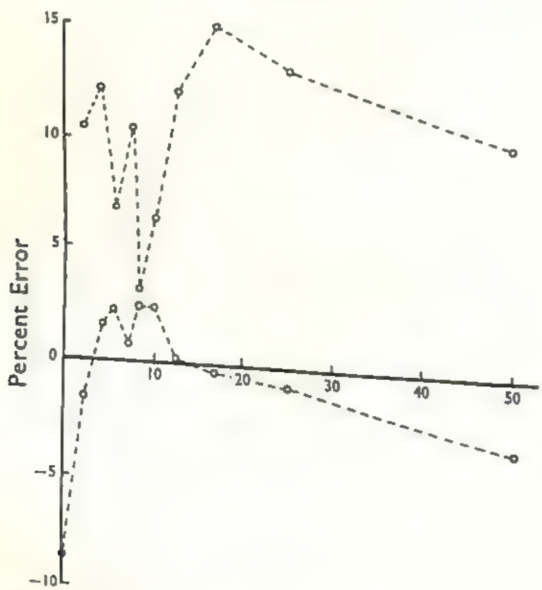
CONTRAST EFFECTS IN DIVIDING DISTANCE AND THEIR SIGNIFICANCES

Figure	1	3	5	7	9	11	13	17	24	49
Mean contrast mm.	5·35	3·34	2·49	1·51	0·63	0·26	0·74	0·38	0·49	0·21
Mean contrast per cent.	10·7	13·4	14·9	12·1	6·3	3·1	10·4	6·8	12·2	10·5
S.D. mm.	5·98	2·91	2·52	2·87	1·69	1·44	1·31	1·25	0·96	0·79
S.D. per cent.	12·0	11·6	15·1	23·0	16·9	17·3	18·3	22·5	24·0	39·6
t ($d.f.$ = 44)	3·00	3·84	3·32	1·76	1·25	0·63	1·89	1·02	1·71	0·89

$P(t \geq 1{\cdot}68) = 0{\cdot}05$; $P(t \geq 2{\cdot}41) = 0{\cdot}01$; $P(t \geq 3{\cdot}28) = 0{\cdot}001$

In Table VI the percentage mean contrast should be the same as the total illusion predictable on the basis of the contrast effects (mean contrast of one interval × the number of intervals), on the assumption that the contrast effects are additive. Figure 4 shows that the prediction concerning the correspondence of the two types of error in question is not borne out. Perhaps this is due in part to some factors other than properties of contrast or chance. Yet, the curve corresponding to the estimated error has the same form as the curve based on empirical results up to the interval of 5·5 per cent. The product moment correlation between the estimated and the empirical values is −0·56. The negative value of the correlation is possibly

FIGURE 4



Percentage contrast effect and the illusion of divided distance predicted from the contrast effect (the upper points; $n = 22$) and the percentage illusion of divided distance found experimentally (the lower points; $n = 24$). The horizontal axis shows the ratio between the dividing distance and the whole distance.

explained by referring to overrepresentation of small distances in the correlation coefficient and large random variation with the small values. In any case the results suggest that the contrast effects are not additive, because an excessive illusion is predicted on the basis of the contrast effects.

Figure 4 shows clearly a curvilinear relationship between the number of intervals and the amount of illusion. In the totally black Figure F the error of estimation

reached the value $-8 \cdot 67$ mm., which differs from the mean error of estimation in stimulus figure O ($-3 \cdot 42$ mm.) at a highly significant level ($t = 4 \cdot 23$, $d.f. = 23$). Stimulus F was thus underestimated. All the other stimulus figures including the bisected distances were overestimated. The results cannot easily be explained by referring to constancy.

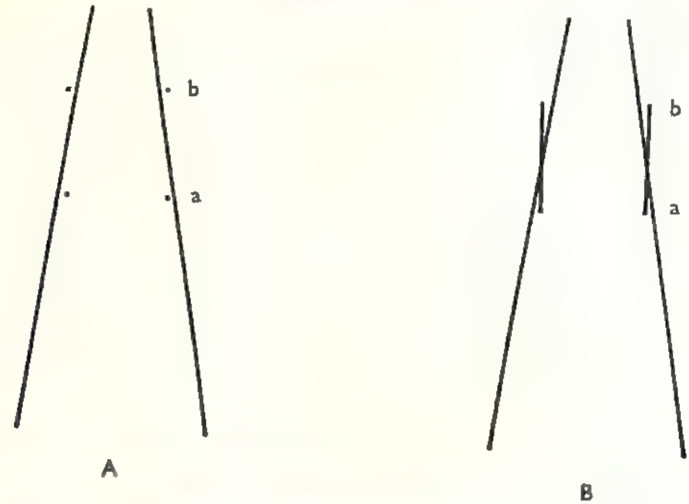
Discussion

The results which were obtained are in accordance with the predictions made on the basis of the contrast explanation. The predictions based on the constancy theory turned out to be unsatisfactory in the figures used here. The figure in Experiment 1 has a tendency to bring about secondary scaling. Explanation of the results by reference to primary scaling could make them more intelligible; this explanation would, however, clearly be of an *ad hoc* type. Secondary scaling is not likely to occur in the figure of divided distance. But how could it be explained by reference to primary scaling? It is clear that the constancy theory in its present form does not give adequate initial conditions for explanation: it is hard to decide when the constancy scaling takes place, and when one should speak of primary vs. secondary scaling. On the other hand, there is evidence enough that the constancy theory will not yield a sufficient explanation of all geometric illusions and of the regularities which have been found even with adjustments to the theory. This may be supported also by Jahoda's finding that among tribes in Ghana there were not differences in Müller–Lyer illusion in the direction predictable from the constancy theory (Jahoda, 1966).

In his discussion of theories of geometric illusions Gregory does not mention the contrast explanation (Gregory, 1966, pp. 141–145). We agree with Gregory that all the other general theories except the constancy theory may be rejected, if we do not count the contrast explanation a theory. However, contrast and confluxion are concepts worthy of consideration, since the use of them, as we have seen, can provide successful predictions. No doubt there is conceptual vagueness in the contrast explanation, and the use of opposite principles of explanation may be dangerous. But, should we not base our explanation on opposite principles, if opposing mechanisms exist which would account for the perceptual processes? As a matter of fact, the vagueness of concepts which could be applied in a contrast theory if developed would not be greater than what we encounter in the constancy theory. Both contrast and confluxion can be interpreted to represent biologically meaningful forms of stimulus generalization. The generality and biological adequacy of contrast and confluxion are almost comparable to that of constancy.

The significance of Experiment 2 lies in the fact that it makes the explanation of illusions occurring in divided figures by reference to the contrast phenomenon plausible. The contrast explanation is, however, far from being sufficient to explain all illusions. For example, explanation of the illusions of direction by reference to contrast and confluxion is not successful. According to Morinaga a "paradox of displacement" appears in Figure 5: the distance between the points *b* looks smaller than the distance between the points marked *a* in Figure 5A, but in Figure 5B the distance between the end points *b* of the vertical lines seems to be longer than the distance between the other end points *a* (Oyama, 1960). The conclusion Morinaga draws is that "the illusions in direction are independent of the illusions in distance or in location." The direction of the illusion in Figure 5A can be explained by referring to confluxion, but the same explanation cannot be applied to Figure 5B. As Ebbinghaus has remarked (Ebbinghaus, 1913, p. 74) Lipps's figure (Fig. 1)

FIGURE 5



is a fragment of Zöllner's figure; nevertheless, the illusion in Lipps's figure is in the opposite direction. This further supports the suggestion that contrast-confluxion and the overestimation of acute angles are two separate phenomena. Most probably the overestimation of acute angles is a phenomenon contributing to the occurrence of illusions. "A general rule which applies to all cases of figural distortion on a background field is that acute angles are perceptually enlarged" (Wallace, 1966).

Overestimation of acute angles may be classified as an example of a tendency to form constancy. Gregory's experiments with the Necker cube also strongly suggest that size constancy should be accepted as a factor producing illusion. This is in accordance with the finding, which has not been mentioned by Gregory, that in the Necker cube drawn on paper, the *edge* which is located further back looks bigger (Gregory, 1966, p. 12, Fig. 1.4). According to Gregory "although a face alternates with another in depth, they do not change in size" (Ibid., p. 156). On the other hand, Tausch's experiments give convincing evidence for rejection of an explanation of the original Ponzo illusion by reference only to contrast and confluxion (Tausch, 1954). Effects of inappropriate size constancy scaling may be further seen in exponents of certain psychophysical power functions. Using the method of ratio estimation, Ekman and Junge (1961) obtained exponents ranging from 0·74 to 0·79 for the volume of drawn cubes and spheres. For the volume of solid cubes they obtained the exponent value 1·01. It still is not clear why constancy scaling did not appear in the estimation of drawn lines, squares, and circles. The exponents for these stimuli varied from 0·92 to 1·11.

It is probable that not all illusions are explicable in terms of only one strictly confined principle. Presumably constancy, contrast and confluxion are all phenomena which contribute to the formation of illusions. In most figures constancy, contrast and confluxion, operate inseparably and in the same direction. In the development of a theory of geometric illusions these factors, all of which are biologically meaningful, should be taken into account. This kind of theory might suffice to explain geometric illusions. However, this does not imply that there are not other biologically meaningful mechanisms of perception which may produce illusions in inappropriate conditions.

This investigation was carried out by a grant from the Board of the National Research Councils.

References

Cook, H. D. (1909). Die taktile Schätzung von ausgefüllten und leeren Strecken. *Arch. ges. Psychol.*, **16**, 418–548.

Cymbalistyj, B. J. (1949). Grössenangleichung und Grössenkontrast in den geometrisch-optischen Täuschungen. *Psychol. Forsch.*, **23**, 115–84.

Ebbinghaus, H. (1913). *Grundzüge der Psychologie*. Vol. II. Leipzig: Veit & Comp.

Ekman, G., and Junge, K. (1961). Psychophysical relations in visual perception of length, area and volume. *Scand. J. Psychol.*, **2**, 1–10.

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, **199**, 678–80.

Gregory, R. L. (1965). Seeing in depth. *Nature*, **207**, 16–7.

Gregory, R. L. (1966). *Eye and Brain: the psychology of seeing*. London: World University Library.

Jahoda, G. (1966). Geometric illusions and environment: a study in Ghana. *Brit. J. Psychol.*, **57**, 193–9.

Kristof, W. (1961). Über die Einordnung geometrisch-optischer Täuschungen in die Gesetzmässigkeiten der visuellen Wahrnehmung. *Arch. ges. Psychol.*, **113**, 1–48 and 127–50.

Oyama, T. (1960). Japanese studies on the so-called geometrical-optical illusions. *Psychologia*, **3**, 7–20.

Révész, G. (1934). System der optischen und haptischen Raumtäuschungen. *Z. Psychol.*, **131**, 296–375.

Rudel, R. G., and Teuber, H.-L. (1963). Decrement of visual and haptic Müller–Lyer illusion on repeated trials: a study of crossmodal transfer. *Quart. J. exp. Psychol.*, **15**, 125–31.

Tausch, R. (1954). Optische Täuschungen als artifizielle Effekte der Gestaltungsprozesse von Grössen- und Formenkonstanz in der natürlichen Raumwahrnehmung. *Psychol. Forsch.*, **24**, 299–348.

Wallace, G. K. (1966). Optical illusions. *Nature*, **209**, 327–8.

Virsu, V. (1966). Note on the systematic error of estimation as a function of stimulus magnitude. *Scand. J. Psychol.*, **7**, 76–80.

Manuscript received 10th March, 1967.

REVERSAL OF THE MÜLLER-LYER ILLUSION WITH CHANGES IN THE LENGTH OF THE INTER-FINS LINE

BY

BRIAN J. FELLOWS

From the Department of Liberal Studies, Portsmouth College of Technology

Subjects were asked to judge the relative lengths of lines placed between the ingoing and the outgoing fins of a Müller–Lyer figure. It was found that with the gap between the fins constant at 160 mm. there was a reversal of the normal Müller–Lyer illusion when the inter-fins line was between 40 and 100 mm. long. This effect was maximal when the line was 80 mm. The normal illusion returned when the line was increased in length to 120 mm. These findings do not support Gregory's inappropriate constancy scaling theory. They seem to suggest the operation of two distinct illusory forces in the Müller–Lyer situation. The reverse illusion may be due to the greater expansion effect upon the line between the ingoing fins, caused by the greater "enclosing" nature of these fins.

Introduction

Gregory (1963, 1966) has proposed that the Müller–Lyer, and certain other, illusions are the result of inappropriate size constancy scaling. Since the illusion figures can be thought of as flat projections of common three-dimensional situations, Gregory argued that they may be processed by the visual system as if they were three-dimensional. More specifically, it is argued that our perception of the flat illusion figures may be modified by the constancy scaling mechanism which is known to affect the perception of size in three-dimensional situations (Thouless, 1931). So, in the Müller–Lyer figure, the outgoing fins may correspond to the pattern made by the inside corner of a room; the vertical line will be relatively distant and, hence, phenomenally expanded. On the other hand, the ingoing fins may represent the outside corner of a building; here, the vertical line will be relatively near and, hence, phenomenally contracted.

Assuming that the constancy scaling mechanism can account both for expansion and for contraction (which is not clear), one might predict from this theory that any object placed between the fins of the Müller–Lyer figure would be subject to the same pressures as is the connecting line in the classical figure. Thus, we might expect an object placed between the outgoing fins to be expanded and one placed between the ingoing fins to be contracted. Casual observation does not support this expectation. In fact, as may be seen from Figure 1, the opposite seems to occur. In this figure, the line between the ingoing fins appears longer than the line between the outgoing fins though both are physically of the same length.

The present experiment was designed to investigate this unexpected observation in more detail. To what extent does the reversal of the Müller–Lyer illusion vary with the length of the inter-fins line? Is there a point where the effect is maximal? Is there a balance point where the classical Müller–Lyer effect and its reverse cancel each other out to produce perfect judgements?

Method

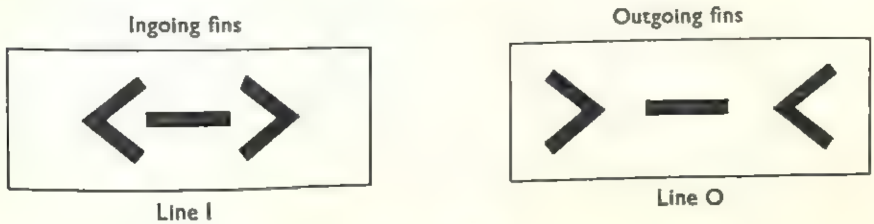
Apparatus

On each of two white cards 13 × 5 in., two pairs of fins were drawn in black Indian ink. These are illustrated in Figure 1. Measured from the inside corner, the length of each fin was 50 mm. The fins were 10 mm. wide and the inside angle was 80°. The

gap between the outgoing fins (outside corners) and the ingoing fins (inside corners) was 1·9 mm.

For the inter-fins lines, seven sets of six black rectangular cards were prepared, all 10 mm. wide and approximately 0·5 mm. thick. Each set contained one standard length and five comparison lengths of measurements as shown in Table I.

FIGURE 1



M–L Cards.

Subjects

Sixteen first year college students, all studying psychology as part of a General Degree, were employed as subjects. All subjects were familiar with the Müller–Lyer illusion and well experienced at making visual judgements of length in experimental situations.

Procedure

The experiment was conducted in a well lit room. Experimenter and subject stood on opposite sides of a low table (2 ft. 6 in.). On the table between them stood a 2 ft. high screen. There was a ½ in. gap between the bottom of the screen and the table. The experimenter placed the appropriate rectangles in position on the M–L cards, midway between the fins, and then pushed the two cards along the table, and through the gap under the screen for the subject to look at them and make his judgement.

The subject was instructed to stand close to the table and to look down at the two rectangles, and to indicate, by raising either his left, or his right arm, which rectangle looked the longer. It was emphasized that the judgement should be based upon how the rectangles actually looked and not upon any preconceptions the subject may have about the effects of the fins, or any other "rational" factors. It was stressed that there were no right or wrong judgements. No feedback about the judgements was given. Pilot work indicated that an "Equal" category was demanded by the subject. It was thought that the subjects employed in this experiment were psychologically sophisticated enough to use the "Equal" judgement appropriately and, hence, avoid the possible methodological drawbacks of its inclusion (the results showed this assumption to be justified). The subject was not informed that the gaps between the fins on the two cards were equal. This was done to avoid his judgement of the relative lengths of the rectangles being affected by the relative sizes of the gaps between the ends of the rectangles and the fins. Immediately the subject responded, the cards were withdrawn and rearranged for the next trial. The interval between a response and the next presentation was approximately 15 sec.

There were seven experimental sessions, with 10 trials in each session. A different standard length was employed in each session. Each comparison length was presented twice per session; once in the first half of the session and once in the second half; once between the ingoing fins and once between the outgoing fins; and once to the left of the standard and once to the right. The positions of the two M–L cards were alternated from one trial to the next. The position occupied by the longer rectangle (*a*) between the ingoing or the outgoing fins, and (*b*) to the left or to the right of the subject was randomly varied within each session. The sequence of standard lengths used from one session to the next was randomly varied between subjects. There were two experimental periods. In the first, four sessions were completed and in the second, one week later, the remaining three sessions. Otherwise, approximately 20 min. elapsed between sessions.

Results

Table I shows the judgements of the relative lengths of the inter-fins lines for each standard length. It indicates the total number of subjects judging each comparison line to be either longer than, equal to, or shorter than the corresponding

TABLE I

Judgements of the Relative Lengths of the Inter-Fins Lines

Line I					Standard (mm.)	*Line O*				
Comparison (mm.)						*Comparison (mm.)*				
22	*21*	*20*	*19*	*18*	*20*	*18*	*19*	*20*	*21*	*22*
15	8	5	1	1	Longer	0	1	6	12	15
1	7	5	2	1	Equal	1	2	5	3	1
0	1	6	13	14	Shorter	15	13	5	1	0
42	*41*	*40*	*39*	*38*	*40*	*38*	*39*	*40*	*41*	*42*
13	11	8	3	1	Longer	0	0	2	5	8
3	4	6	4	5	Equal	0	1	6	5	3
0	1	2	9	10	Shorter	16	15	8	6	5
64	*62*	*60*	*58*	*56*	*60*	*56*	*58*	*60*	*62*	*64*
16	13	9·5	5	2	Longer	0	0	2	5	11
0	3	4·5	5	4	Equal	0	2	4·5	5	4
0	0	2	6	10	Shorter	16	14	9·5	6	1
84	*82*	*80*	*78*	*76*	*80*	*76*	*78*	*80*	*82*	*84*
16	16	11	8	1	Longer	0	0	1·5	5	5
0	0	3·5	6	4	Equal	0	3	3·5	4	6
0	0	1·5	2	11	Shorter	16	13	11	7	5
104	*102*	*100*	*98*	*96*	*100*	*96*	*98*	*100*	*102*	*104*
16	13	7·5	4	4	Longer	2	1	5·5	5	9
0	1	3	5	4	Equal	2	3	3	4	4
0	2	5·5	7	8	Shorter	12	12	7·5	7	3
124	*122*	*120*	*118*	*116*	*120*	*116*	*118*	*120*	*122*	*124*
13	5	4·5	3	1	Longer	3	2	8	12	12
2	4	3·5	6	2	Equal	3	4	3·5	3	3
1	7	8	7	13	Shorter	10	10	4·5	1	1
144	*142*	*140*	*138*	*136*	*140*	*136*	*138*	*140*	*142*	*144*
11	2	1·5	0	0	Longer	2	8	12·5	15	16
3	3	2	2	0	Equal	4	4	2	1	0
2	11	12·5	14	16	Shorter	10	4	1·5	0	0

standard line when the comparison line is situated either between the ingoing fins (when it is referred to as *Line I*), or between the outgoing fins (when it is referred to as *Line O*). The "Shorter" judgements are derived from the subjects' "Longer" and "Equal" judgements. Since the two lines of equal length in each set were used

interchangeably as standard and comparison, the judgements with respect to them have been summed and halved to give the figures shown in the table.

The effects of the length of the standard line upon the judgement of the relative sizes of Line I and Line O can be seen more clearly in Table II.

TABLE II

"LONGER" AND "EQUAL" JUDGEMENTS FOR EACH PHYSICAL DIFFERENCE BETWEEN THE TWO INTER-FINS LINES

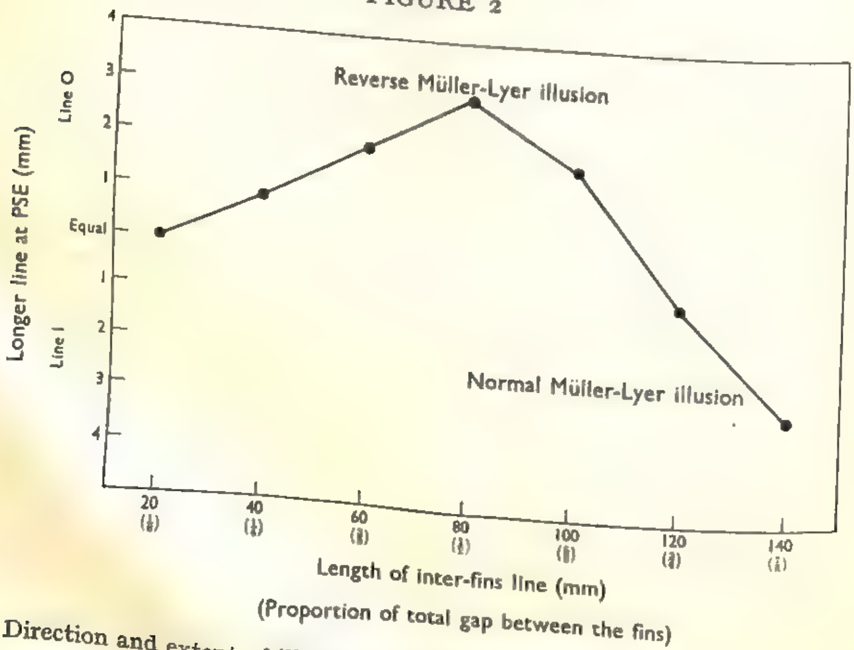
	Longer line					*Standard*
	Line I		*Equal*	*Line O*		
Judgement	*2 mm.*	*1 mm.*	*0*	*1 mm.*	*2 mm.*	*20 mm.*
Line I ..	30	21	10	2	1	
Equal ..	2	9	10	5	2	
Line O ..	0	2	12	25	29	
	2 mm.	*1 mm.*	*0*	*1 mm.*	*2 mm.*	*40 mm.*
Line I ..	29	26	16	9	6	
Equal ..	3	5	12	9	8	
Line O ..	0	1	4	14	18	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*	*60 mm.*
Line I ..	32	27	19	11	3	
Equal ..	0	5	9	10	8	
Line O ..	0	0	4	11	21	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*	*80 mm.*
Line I ..	32	29	22	15	6	
Equal ..	0	3	7	10	10	
Line O ..	0	0	3	7	16	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*	*100 mm.*
Line I ..	28	25	15	11	7	
Equal ..	2	4	6	9	8	
Line O ..	2	3	11	12	17	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*	*120 mm.*
Line I ..	23	15	9	4	2	
Equal ..	5	8	7	9	5	
Line O ..	4	9	16	19	25	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*	*140 mm.*
Line I ..	21	6	3	0	0	
Equal ..	7	7	4	3	0	
Line O ..	4	19	25	29	32	

In this table the number of "Longer" judgements for each line and the number of "Equal" judgements are shown for each physical difference between the two lines (irrespective of which is standard and which is comparison). For example, with a

standard length of 80 mm., when Line I was 4 mm. longer than Line O (i.e., when Line I = 84 mm. and Line O = 80 mm., and when Line I = 80 mm. and Line O = 76 mm.) Line I was always judged to be the longer; when Line I was 2 mm. longer it received 29 "Longer" judgements and was judged to equal to Line O three times; when the two lines were equal Line I received 22 "Longer" judgements, Line O received three and seven judgements made them equal. From Table II the trend of the judgements can be most easily seen by inspecting the figures in the "Equal" column. It can be seen that when a standard of 20 mm. was employed there was very little discrepancy between the judgement and the actual lengths of the inter-fins lines. However, as the standard increased in length up to 80 mm. the tendency was for Line I to be judged the longer. With further increases in the length of the standard line this tendency was reversed until at 140 mm. Line O was more frequently judged to be the longer.

On the basis of the data in Table II a point of subjective equality (PSE) was determined for the two lines at each standard length. The probabilities of the "Longer" judgements were converted into z scores and plotted against each difference in the lengths of the inter-fins lines. Straight-line graphs for the "Line I Longer" and "Line O Longer" points were fitted using the averaged z scores technique (Woodworth and Schlosberg, 1954), thus effectively utilizing all the data. The PSE's were determined by dropping lines from where the graphs crossed. Figure 2 shows the required difference for PSE at the various standard lengths. With a standard of 20 mm. (one-eighth of the gap between the fins) there is no illusion. Increasing the standard to 40 mm. (one-quarter of the gap between the fins) produces a reversal of the normal Müller–Lyer illusion; here, the line between the outgoing fins (Line O) needs to be 1 mm. longer than the line between the ingoing fins (Line I) for them to appear equal (PSE). The tendency to underestimate Line O as compared with Line I increases when the standard length is increased to 60 mm. (three-eighths of the gap between the fins); here, Line O needs to be 1·9 mm. longer than Lines

FIGURE 2



Direction and extent of illusion for different lengths of inter-fins line.

I for PSE. The reverse Müller–Lyer illusion appears to be strongest when the inter-fins line is 80 mm. (one-half of the gap between the fins); here, Line O needs to be 2·8 mm. longer than Line I for subjective equality. Increasing the length of the inter-fins line to 100 mm (five-eights of the gap between the fins) reduces the required difference for PSE to 1·6 mm. Increasing the standard length further to 120 mm. (three-quarters of the gap between the fins) reverses the effect, thus restoring the normal Müller–Lyer illusion; here, Line I needs to be 1 mm. longer than Line O for subjective equality. The strength of the Müller–Lyer effect is increased when the length of the inter-fins line is raised to 140 mm. (seven-eighths of the gap between the fins); now, PSE occurs when Line I is 2·9 mm. longer than Line O. The balance point, where the normal Müller–Lyer and the reverse Müller–Lyer effects cancel each other out, seems to occur when the standard is about 110 mm.

Discussion

The present experiment has shown, firstly, that the normal Müller–Lyer illusion is reversed when the inter-fins line is somewhat shorter than the total gap between the fins and, secondly, that the effect varies in a regular manner with changes in the length of the inter-fins line. The reverse effect seems to be maximal when the line is about half the size of the gap. Increasing the length of the inter-fins line produces a progressive swing back to the normal Müller–Lyer effect. These findings fail to support Gregory's inappropriate constancy scaling theory which would predict that the shorter lines would be subject to the same effects as the complete shaft in the Müller–Lyer figure.

There appear to be two possible ways of explaining these findings. Firstly, one may argue that the same process is involved as in the normal Müller–Lyer illusion, but, for some reason, the effect is reversed when the length of the line is reduced relative to the gap between the fins. A second, and possibly more fruitful hypothesis, is to postulate the operation of some other process as being responsible for the reversal effect. Thus, the process involved in the normal Müller–Lyer illusion works most effectively when the inter-fins line is complete, but as this line is shortened the second process gradually exerts its influence to reverse the effect. A likely candidate for this latter role is that process which is responsible for making objects in enclosed spaces look larger than the same sized objects in open spaces. The ingoing fins clearly have a greater "enclosing" effect than the outgoing fins and, therefore, should be expected to expand the short interposed line far more.

One prediction from this hypothesis would be that since the enclosure effect produces expansion of objects in all directions the line between the ingoing fins should not only look longer than the line between the outgoing fins, but also should look fatter. This was informally confirmed in the present experiment from the spontaneous comments of subjects to this effect.

A second prediction concerns the enclosing effect of the outgoing fins. It may be argued that though this effect is much less strong than that of the ingoing fins, nevertheless, space on the card is restricted by the fins and, hence, there should be some expansion effect relative to a plain card with no fins. This prediction is, however, complicated by the fact that the outgoing fins do appear to have an expansion effect on the whole space in between and, so, might equally well be expected to produce a slight contraction effect relative to that of a plain card. To test these predictions the 16 subjects employed in the main experiment were recalled for one further session in which the 80 mm. set of inter-fins lines was used. The comparisons to be made were between unenclosed lines on a plain card the same size as the M–L

cards (Line X) and lines positioned (*a*) between the ingoing fins (Line I) and (*b*) between the outgoing fins (Line O). The results are shown in Table III and set out

TABLE III

"Longer" and "Equal" Judgements in Comparing an Unenclosed Line (Line X) with (*a*) Line I and (*b*) Line O. Standard Length = 80 mm.

(*a*) Judgement	*Longer line*				
	Line I		*Equal*	*Line X*	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*
Line I	32	27	23	10	3
Equal ..	0	4	7	10	10
Line X ..	0	1	2	12	19
(*b*) Judgement	*Line O*		*Equal*	*Line X*	
	4 mm.	*2 mm.*	*0*	*2 mm.*	*4 mm.*
Line O.. ..	27	24	8	3	2
Equal	3	5	12	8	3
Line X ..	2	3	12	21	27

in the same manner as the data in Table II. The similarity of the distribution of judgements under (*a*) to that of the 80 mm. standard in Table II confirms that the main factor in the reverse Müller–Lyer illusion is the expansion effect of the ingoing fins. For PSE Line X needs to be 2·4 mm. longer than Line I as compared with 2·8 mm. for Line O against Line I. Also, as can be seen under (*b*) in Table III there is no significant tendency in either direction when comparing Line X with Line O. Here, PSE occurs with Line O 0·2 mm. longer than line X, thus indicating no particular expansion or contraction effects produced by the outgoing fins.

References

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, **199**, 678–80.

Gregory, R. L. (1966). *Eye and Brain: The Psychology of Seeing*. London: Weidenfeld & Nicolson.

Thouless, R. H. (1931). Phenomenal regression to the "real" object. I. *Brit. J. Psychol.*, **21**, 339–59.

Woodworth, R. S., and Schlosberg, H. (1954). *Experimental Psychology*. London: Methuen.

Manuscript received 14th March, 1967.

INTERMANUAL TRANSFER OF PRACTICE DECREMENTS WITH A HAPTIC ILLUSION

BY

R. OVER

From the Psychology Department, University of Otago

Sixteen subjects were given 8 pre-practice, 80 practice, and 8 post-practice trials with the inverted T illusion figure. Judgements were made haptically with one hand being used on pre- and post-practice trials and the other hand on practice trials. Mean illusion diminished in magnitude over the practice trials. Inter-manual transfer of the practice decrement was found, but it was partial rather than complete. The illusion after practice was less than before practice, but it was greater than the error found on the last block of practice trials. The bearing of these results on the different theories of illusions is considered.

Introduction

Köhler and Fishbank (1950) have treated geometrical illusions as instances of contour displacement dependent on cortical satiation. They have claimed that extent of satiation, and thus the amount of illusion obtained, is modified during prolonged inspection, or over repeated trials. Data showing crossmodal transfer of practice decrements (Rudel and Teuber, 1963) fail, however, to support this analysis of the practice decrements that are found with most illusion figures. A haptic illusion of the same magnitude as the visual illusion is found when a blindfolded subject moves his finger over the Müller–Lyer figure. Decrements in illusion obtained over repeated visual trials transfer to haptic judgement, and vice versa. The satiation theory is unable to account for these data in that it postulates that satiation is restricted to the sensory system stimulated during inspection.

The present experiment provides a further test of the satiation theory. In their account of kinaesthetic after-effects Köhler and Dinnerstein (1947) have claimed, mainly because there is a contralateral projection within the tactile and kinaesthetic sensory systems, that satiation developed through kinaesthetic inspection is restricted to the hand used during inspection. Kinaesthetic after-effects are in fact generally measured by presenting a bar of given width to one hand during the inspection period, and a bar of different width (the standard stimulus) to that hand and a bar of variable width (the comparison stimulus) to the other hand when obtaining post-inspection measures. Data obtained in a situation in which subjects made absolute rather than relative judgements (Singer and Day, 1966) indicate that kinaesthetic after-effects do not transfer intermanually. From the satiation theory it would thus be predicted that practice decrements obtained with a haptic illusion when inspection has involved only one hand would not transfer when judgements are subsequently made with the other hand. Data on the crossmodal transfer of practice decrements (Rudel and Teuber, 1963) suggest, however, that intermanual transfer will be obtained.

Method

Subjects. The 16 experimental subjects and 8 control subjects were undergraduate students. They were previously unacquainted with the task and had no knowledge of the aim of the experiment.

Apparatus. The inverted T figure was studied; equivalent visual and haptic illusions have been found with this figure (Révész, 1953). The figure was constructed of wire of 0·078 in. diameter, and was supported by a wooden frame to which a tapemeasure was attached. The bisected line was 10 cm. long and the bisecting line 20 cm. long. A movable marker could be set by the subject at any position along the bisecting line.

Procedure. The subject was blindfolded during testing. The figure was presented with the bisected line horizontal and the bisecting line extended away from the subject, and was within easy reach of either hand. Each experimental subject made 8 pre-practice settings with one hand, 80 practice settings with the other hand, and then 8 post-practice settings with the hand used initially. No rest periods were given and most subjects required about 1 hr. to complete all settings. Half of the subjects made the practice settings with the right hand and half with the left hand. In making a setting the subject was required to adjust the movable marker so that the distance between the point of bisection and the marker felt to be the same length as the bisected line. The subject was permitted to touch the figure only with the index finger of the designated hand; he was allowed to trace over all parts of the figure as often as he wished before completing his setting. For half of the subjects the marker was at the point of bisection at the start of each trial, and for the other half the marker was at the far end of the bisecting line. No information was given to subjects about the accuracy of their settings.

It is possible that differences may be found between post-practice and pre-practice settings made by the experimental subjects simply as a function of the time elapsing between the settings rather than as a function of the interpolated practice. Control subjects were tested to assess this possibility. Each control subject was given two sets of 8 trials using one hand. The two sets of trials were separated by 1 hr. Half of the subjects used the right hand and half the left, and half of the subjects made settings from each starting position.

Results

Mean measures of illusion were obtained for the experimental group on the pre-practice trials, at different stages during practice, and on post-practice trials. From a test of trend it was found that there were significant differences in mean illusion over all settings made by the experimental subjects ($F = 9{\cdot}67$, *d.f.* 11, 154, $p < 0{\cdot}01$). Mean illusion was independent of the starting position of the movable marker ($F = 1{\cdot}42$, *d.f.* 1, 14, $p > 0{\cdot}05$), and the trials × starting position interaction was not significant ($F = 0{\cdot}88$, *d.f.* 11, 154, $p > 0{\cdot}05$). Further comparisons were made by t tests or by t tests for correlated means, as appropriate.

The mean illusion on practice trials 73–80 was significantly less than that obtained on practice trials 1–8 ($t = 2{\cdot}89$, *d.f.* 15, $p < 0{\cdot}02$); an intramanual practice decrement has thus been demonstrated. The mean post-practice illusion found for experimental subjects was significantly less their mean pre-practice illusion ($t = 2{\cdot}51$, *d.f.* 15, $p < 0{\cdot}05$). The two sets of judgements made by the control subjects, however, were not significantly different ($t = 0{\cdot}42$, *d.f.* 7, $p > 0{\cdot}05$). Thus the smaller illusion found after practice, relative to measures taken before practice, for the experimental subjects can be attributed to the interpolated practice with the other hand and not simply to the time elapsing between the two measures. The differences between mean pre-practice and post-practice measures of illusion was not significantly different from that found between practice trials 1–8 and 73–80 measures ($t = 1{\cdot}45$, *d.f.* 15, $p > 0{\cdot}05$). Measures of illusion obtained on pre-practice trials did not differ significantly from mean measures on practice trials 1–8 ($t = 1{\cdot}29$, *d.f.* 15, $p > 0{\cdot}05$). However, mean illusion on post-practice trials was significantly greater than that found on practice trials 73–80 ($t = 2{\cdot}35$, *d.f.* 15, $p < 0{\cdot}05$). The above data indicate that intermanual transfer of the practice decrement has been partial rather than complete. On post-practice trials experimental subjects demonstrated a smaller illusion than would be found if they had had no practice. The obtained illusion

was, however, greater than that found at the end of the practice session during which the other hand had been used to inspect the figure.

The individual data can be briefly outlined. All subjects showed a definite haptic illusion; for each subject the distance between the point of bisection and the position on the bisecting line at which the marker was set was less than the length of the bisected line on each of the first eight settings. Decrements in the magnitude of the illusion were obtained for 14 of the 16 experimental subjects. For 13 of these 14 subjects the mean illusion on post-practice trials was less than the mean illusion on pre-practice trials, and for the other subject there was no change in mean illusion. Mean illusion on post-practice trials was greater than mean illusion on pre-practice trials for both of the experimental subjects who did not show a decrement over the practice trials.

Discussion

Both the present results and data on crossmodal transfer (Rudel and Teuber, 1963) indicate that decrements in the magnitude of an illusion over repeated trials cannot be attributed to processes which are restricted to the sensory system or sub-system stimulated during inspection. The treatment of practice decrements offered by Köhler and Fishbank (1950) is inadequate in that it has postulated that satiation is restricted in this way.

It has commonly been claimed (e.g. Judd, 1902; Day, 1962) that practice decrements develop from attitudinal changes that occur over repeated trials. The subject is said to shift from a wholistic to a partistic basis of judgement and to become less influenced by surrounding contours when judging the spatial properties of a component of the figure. Intermanual transfer of practice decrements should be found in that an attitude shift over repeated trials should not be restricted to the hand used during inspection. It would be expected that the transfer would be complete rather than partial. The major weakness of the attitudinal account of practice decrements is that it offers little specification of the judgemental shift which is supposed to occur. The account will generate testable implications only when such a specification is available.

Piaget (1961) has treated illusions as products of processes by which spatial information is selectively sampled during inspection. He has claimed that some parts of an illusion figure receive more attention than others and are consequently over-estimated in size. Attention is considered to become more evenly distributed over the total figure during repeated trials. If attention is given a central locus, Piaget's treatment becomes similar to the attitudinal account considered above. If, however, attention is specified and measured in terms of peripheral processes such as fixation Piaget's account is readily testable. It implies that illusions are found in all modalities in which inspection involves the sampling of information over time; that is, in visual and haptic space but not in tactual space. Practice decrements would be attributed to modifications in the inspection pattern adopted by the subject. The degree of intermanual transfer obtained would be dependent on the extent to which the inspection pattern adopted by the hand used after practice is similar to that employed by the other hand at the end of the practice trials. Cross-modal transfer of practice decrements can be considered in similar terms.

This research was supported by a grant from the New Zealand University Grants Committee.

References

Day, R. H. (1962). The effects of repeated trials and prolonged fixation on error in the Müller–Lyer figure. *Psychol. Monogr.*, **76**, Whole No. 533.

Judd, C. H. (1902). Practice and its effects on the perception of illusions. *Psychol. Rev.*, **9**, 27–39.

Köhler, W., and Dinnerstein, D. (1947). Figural aftereffects in kinaesthesis. In *Miscellanae Psychologica Albert Michotte*. Louvain: Louvain University.

Köhler, W., and Fishbank, J. (1950). The destruction of the Müller–Lyer illusion in repeated trials. *J. exp. Psychol.*, **40**, 267–81, 398–410.

Piaget, J. (1961). *Les Mécanismes Perceptifs*. Paris: Presses Universitaires de France.

Révész, G. (1953). Lassen sich die bakannten geometrischoptischen Tauschungen auch in haptischen Gebiet nachweisen? *J. Psychol. Psychotherapie*, **1**, 464–78.

Rudel, R. G., and Teuber, H.-L. (1963). Decrement of visual and haptic Müller–Lyer illusion on repeated trials: A study of crossmodal transfer. *Quart. J. exp. Psychol.*, **15**, 125–31.

Singer, G., and Day, R. H. (1966). Interlimb and interjoint transfer of a kinesthetic spatial aftereffect. *J. exp. Psychol.*, **71**, 109–14.

Manuscript received 27th January, 1967.

COMMENTS ON THE INAPPROPRIATE CONSTANCY SCALING THEORY OF THE ILLUSIONS AND ITS IMPLICATIONS

BY

R. L. GREGORY

From the Psychological Laboratory, Cambridge

In the literature of perception no clear distinction is generally made between: (i) the *processes* leading to perceptual correction for changes in retinal image size with object distance, and (ii) the final *result*—near-invariance of perceptual size with object distance. Thus, J. J. Gibson is not at all concerned with brain mechanisms giving constancy, but regards it as determined by the surrounding pattern of the prevailing external world and to be fully describable in terms of texture gradients and the other pattern features he has done so much to investigate. But we must face the fact that the monocular depth cues are open to many interpretations, and it is difficult to believe that perception is wholly determined by them or can be understood without reference to brain processes involving probabilities. There are two particular situations which drive us away from Gibson's position: (1) depth-ambiguous figures, such as luminous Necker cubes, which change in shape following Emmert's Law with each reversal though the retinal pattern remains constant; (2) the limiting case of a single object having no visible background to determine its size or distance. Gibson does not consider the first of these, and says of the second that size and distance are indeterminate under these conditions. But the first shows dramatic and repeatable effects, not to be dismissed lightly, while for the second Gibson is forced by his position into statements empirically incorrect. He is forced to say that the moon, or an after-image viewed in darkness, have *no* perceptual size or distance. But they have a quite clear size and distance, which can indeed be measured. They may fluctuate, and may be wildly in error, but we never see an object or after image having *no* perceptual size or distance. Luminous figures viewed in darkness reveal that we accept hypothetical sizes and distances. Depth-ambiguous luminous figures change in step-functions, revealing alternative perceptual hypotheses—and perceptual constancy follows each hypothesis of distance. It is thus clear that constancy is not always tied to the external visual pattern. A "stimulus" theory cannot be adequate, because it cannot describe the ambiguous figure case, or the limiting case of zero context information.

Gibson almost completely dismisses the distortion illusions, at times saying they do not exist. They are difficult to reconcile with a view of perception which regards percepts as determined by the pattern of the external world, without reference to brain processes which may be upset in certain situations. This leads us back to the distinction between (i) the processes leading to constancy and (ii) the size-distance invariance itself. We term the *processes* "constancy size scaling" (or for short "constancy scaling"), implying active brain processes, producing the size-distance invariance, to give *constancy*.

Considering now the scaling processes: the theory essentially involves two very different kinds of scaling, the first being tied to typical prevailing depth information, while the second is tied to the current perceptual hypothesis. Up to now I have called these "primary" and "secondary" scaling respectively, but it may be convenient

to identify them by more explicit names. I suggest: "*Depth cue scaling*" and "*hypothesis scaling*" for the processes giving constancy. Illusions occur when *either* of the scaling systems is set inappropriately to the prevailing three-dimensional external world. The standard visual illusions are to be attributed to *depth cue scaling* being set by perspective depth cues inappropriate to the flat figures. Other depth cues can give distortion when set inappropriately, such as incorrect convergence of the eyes through deviating prisms, or the "wall paper" effect when horizontally displaced repeated patterns are fused with resulting error in size. It is a matter for experiment to determine whether all depth cues can give distortion when set inappropriately to prevailing reality. From such experiments we can learn a great deal about the brain's scaling mechanisms.

Hypothesis scaling is very different. It is not always closely related to the prevailing retinal information, as we see from the depth-ambiguous figures, which change in size and shape with each perceptual reversal. Hypothesis scaling is subject to modification by the observer's "set," by his pre suppositions, and is more variable than is depth cue scaling. Perceptual hypotheses can be wildly wrong but they have the great biological value, in an uncertain world, of filling in gaps between available items of information—allowing usually appropriate behaviour to be maintained though there is no certain sensory information to control it. The Favourite cannot always win, but an Outsider does not generally lead to disaster. Illusions can normally be tolerated, but highly unusual conditions, such as space travel, may defeat the system when perceptual hypotheses run away from sensory control.

It is in the light of these general considerations that I now turn to consider the points made by Zanforlin, Virsu, Fellows and Over in this *Journal.*

Dr. Zanforlin, in his first paragraph, says that I attribute the distortions to conflict of depth information from the figure and its background. This is not however quite correct, and the point is important both for understanding and for testing the theory. The relevant fact is that distortions are observed in the illusion figures whether or not they have a visible background. We cannot therefore attribute the distortions to conflict between depth features of the figure and its background. Experiments with luminous figures, having no visible background, show that: (1) the distortions persist though the background be removed; (2) the figures then generally appear to lie in three dimensions (though truly flat), according to their perspective features. They no longer lie in the queer paradoxical depth of pictures with backgrounds—where though depth is suggested the figure is seen as lying flat on its background and in (paradoxical) depth. (3) The distortions are affected by the non-paradoxical depth: apparent size of the figure as a whole, and of each feature, now obeying Emmert's Law. (This is seen most clearly with depth-ambiguous figures such as a luminous Necker cube, the apparently further face always appearing the larger, whichever this may be.) (4) Depth-ambiguous illusion figures show both the residual illusions and size changes following Emmert's law upon reversal. (5) Necker cubes presented with a background do not change appreciably in shape with reversal in (paradoxical) depth.

The theory does not attribute the illusions to conflict between the figure's perspective depth features and its background because the background is not necessary for the distortion. What the theory does assert is that the distortions occur in figures or objects having marked typical perspective features which are at variance with the *true shape* or *orientation* of the figure or object. For example: the converging lines of the Ponzo figure, or the arrow heads of the Müller–Lyer figure mislead the perceptual constancy mechanism because these shapes are generally reliable perspective

guides to the third dimension; but in these flat figures the eyes are misled because the perspective instead of being generated by the usual geometrical shrinking of the image with distance is present in the (flat) object itself.

This takes us to Zanforlin's doubt about my ability to define "depth features," and his complaint that perspective representations can be ambiguous. I would go further and say that perspective is *always* ambiguous, and that it is this ambiguity which presents a major and continual problem to the perceptual system. We should think of this as a problem confronting any conceivable perceptual system allowed only two-dimensional representations of three-dimensional reality. The ambiguity results from the loss of a dimension at the eye, and perceptual theory should start with this basic situation. Ambiguity of depth is reduced as much as possible by the use made of the various "depth cues" (investigated by Helmholtz, Ames, the Gibsons, and others), typical perspective shapes being used by post early Renaissance artists to indicate the third dimension with fair success. Artists are, however handicapped by the texture of their backgrounds, and it is amazing that we see any depth under these conditions and it may require special perceptual learning. It is important to note that a picture has a kind of double reality, and we cannot suppose that constancy scaling can be set correctly both for the picture as a flat object and for the objects it depicts in the quite different picture space. All figures are in this sense "impossible objects," and the amazing thing is how well we handle them. Common sense is a fair guide to what is a "typical" perspective projection, but one could take a set of photographs from normal viewpoints as a check. Obviously a square is not a typical perspective projection of a truncated cone viewed obliquely, and a view through opposite corners of a cube is not typical, or characteristic, either. Perhaps more important is what the perceptual system accepts as reliable perspective depth information. This is an empirical question open to experiment. Zanforlin's objection that we cannot establish depth features by presenting self-luminous figures in the dark (and measuring their effects quantitatively with the technique in which the two eyes are used as a range-finder for plotting monocular visual space) must not go without comment. He objects to the "dubious predictions" of this technique, but surely he is pre-judging the issue without alternative evidence: and the measure of the power of a technique is the surprise of the results it can give. He should be clearer that the results are impossible, or that the technique is subject to serious artifacts, before denouncing it. A result of the technique (and not wholly surprising) is that for two identical shapes, one larger than the other, the larger is generally seen as nearer. It is absurd to argue *a priori* that the perceptual system must or should take a 50–50 per cent. view of the situation—it is a matter for experiment to discover its strategy, and the technique allows us to discover how it deals with such ambiguous situations quantitatively and under any desired conditions. In fact, statistically, large images represent nearer objects and it turns out that the brain accepts this probability as a depth cue. As indeed we know from the many alley experiments (Ittelson and Kilpatrick, 1952).

Perception of depth is not *completely* determined by texture or perspective. In impoverished environments the observer's preconceptions and his "set" can be important. These "internal" factors can be assessed from the depth measurements, and they are important for astronauts and others required to make perceptual judgement under impoverished conditions. The limiting case, of a single luminous object viewed in darkness, has been too-little considered. It is remarkable that the moon is so constant in apparent size and distance when viewed in a clear sky giving no obvious depth information, and it evidently appears much the same size and distance to all observers—though of course we all see it quite wrongly! It seems that

depth information serves to modify "hypotheses," perhaps themselves derived from experience of many earlier situations accepted as relevant to the present conditions. Where there is little or no depth information, the prevailing "perceptual hypothesis" may go unchecked, then size and distance vary wildly, especially in unfamiliar situations. Systematic illusions occur when the depth information is misleading through being atypical. The illusion figures give systematic distortion because they present perspective features which are typical of quite different objects, lying in three dimensions. Distortions are not however limited to flat objects—they are but one case; for any object can be so shaped that it displays misleading depth information which, it seems, can set constancy inappropriately to its true shape and so generate an illusion.

Dr. Virsu's learned paper on contrast and confluxion does not need detailed comments. I find the notion rather vague, and would prefer to push what seems a more precise set of concepts—primary and secondary constancy scaling—to the limit before accepting them as incomplete, but certainly we do not have the evidence to assert that misplaced constancy is the *whole story* of illusions. Virsu could be correct; but if so the perceptual system is "messier" than one might hope. Is there evidence to show this?

B. J. Fellows' experiment—showing that a line placed between the arrow heads of the Müller–Lyer figure but of insufficient length to reach either arrow may show a reversed illusion—strikes me as neat and perhaps important. I am not however clear why he argues with such confidence that: "these findings fail to support Gregory's inappropriate constancy scaling theory which would predict that the shorter lines would be subject to the same effects as the complete shaft of the Müller–Lyer figure." Consider first what is well known about this illusion:

(1) The Müller–Lyer illusion is unusual in that the figure *itself* is distorted, while generally it is imposed lines which are distorted. But here it is the separation of the arrows which is in error whether or not there is an intervening line.

(2) The outward going arrows give expansion, and the inward ones shrinking *with respect to a neutral comparison line*. There can, then, be *expansion* or *shrinking* in the illusions.

(3) The distortion is a change of *separation* between the arrow heads, and not of the *angles* of the figure, which is of theoretical importance.

Now given that the separation between the arrow heads is changed by the usual illusion (and this occurs in the absence of any line joining them) what "should" happen to a short line placed between the heads? If the heads correspond perceptually to the retinal projection of corners, a shorter line could represent, in the case of the outgoing arrows, some object *nearer* the observer than the extreme of the (inside) corner. In the real world this would give a larger retinal image than when placed at the corner; so to give constancy it must be shrunk with respect to the corner—which is what Fellows finds in his experiment. (I would regard this figure as perceptually the same as the Ponzo illusion but, as it were, viewed end on. I would predict that the short line will be measured as perceptually nearer, with the depth-measuring technique*). No doubt there are other possibilities for interpreting this experiment which should be considered (what happens if the arrow heads, or neutral lines, are physically further or nearer the test line?) but to my mind the "enclosing"

* This prediction has since been confirmed when measurements of apparent depth were made in the figures for 16 subjects. This experiment, suggested by Fellows's results, was carried out recently (June, 1967). It is hoped to publish full details later.

effect argument is not attractive—unless by "enclosing" one means the volume of 3-D space represented by the projection, and this is a very different notion.

Regarding the relevance of touch illusions to the illusions of vision (and especially Over's findings, which supplement those of Rudel and Teuber) I offer no comment at this stage. Haptic touch on the Müller–Lyer illusion is confounded by the poor touch acuity of the fingers, which tends to produce a similar effect for figures such as the visual Müller–Lyer though by quite different means. At present the whole question of the relation between touch and vision is too uncertain for us to say how relevant touch experiments are to a theory of visual illusions, but perhaps a close relationship would indicate that the touch and visual spaces are neurally related in the nervous system *after* visual primary constancy scaling has taken place. However this may be, the depth-ambiguity of retinal images implies that non-visual information is needed to give meaning to retinal images in terms of external reality (Gregory, 1967*a*), essential for the development of the individual's visual perception (Gregory and Wallace, 1963), and for the development of the first effective eye-brain systems in evolution (Gregory, 1967*b*). But the relationship between touch and vision in illusion situations is at present largely mysterious, and it is unwise to make any specific statement at the present stage of knowledge.

References

FELLOWS, B. J. (1967). Reversal of the Müller–Lyer illusion with changes in the length on inter-fins lines. *Quart. J. exp. Psychol.*, **19**, 208–14.

GREGORY, R. L. (1967*a*). Will seeing machines have illusions? In COLLINS, N. L. and MICHIE, D. (Ed.), *Machine Intelligence*, 1.

GREGORY, R. L. (1967*b*). Origin of eyes and brains. *Nature*, **213**, 369–72.

GREGORY, R. L., and WALLACE, J. G. (1963). *Recovery from Early Blindedness*. E.P.S. Monogr. No. 2.

ITTELSON, W. H. and KILPATRICK, F. P. (1952). Experiments in perception. *Scient. Amer.*, **185**, 50–5.

OVER, R. (1967). Intermanual transfer of practice decrements with a haptic illusion. *Quart. J. exp. Psychol.*, **19**, 215–18.

VIRSU, V. (1967). Contrast and confluxion as components in geometric illusions. *Quart. J. exp. Psychol.*, **19**, 198–207.

ZANFORLIN, M. (1967). Some observations on Gregory's theory of perceptual illusions. *Quart. J. exp. Psychol.*, **19**, 193–7.

Manuscript received 12th May, 1967.

BINOCULAR RIVALRY AND IMMEDIATE MEMORY

BY

H. SAMPSON and J. B. HORROCKS

From the Department of Psychology, University of Auckland

Three experiments examine features of a simple memory task on which right-handed, right eye dominant subjects have been reported to recall digits projected to the right eye more accurately than those projected simultaneously to the left eye. Superior recall by these subjects of information projected to the right eye was observed only when stimuli projected simultaneously to both eyes were seen as overlapped in the binocular percept. Under *monocular* presentations, accuracy of recall was not related to the eye with which stimuli were viewed. The *binocular* overlap condition has a significance other than that of simply increasing the difficulty of identifying the elements in a visual display for there were no differences in accuracy of recall from each eye when overlapped stimuli were viewed *monocularly*. More accurate recall of right eye information appears to reflect the resolution of a conflict between inputs from each eye. The possible relation of this finding to cerebral dominance is also discussed. Order of recall in these experiments depended mainly on spatial cues provided by the experimental situation.

Introduction

Sampson and Spong (1961*a*, 1961*b*) and Sampson (1964) used a divided viewer to project different stimuli simultaneously to each eye. They found that right-handed, right eye dominant subjects recalled stimuli projected to the right eye more accurately than those projected to the left, and subjects typically reported them in pairs with the stimulus that had been projected to the left eye reported first. These results were interpreted provisionally within the context of a theory of immediate memory and with reference to cerebral dominance. However, there still remains the possibility that further examination of conditions under which these results are obtained could suggest a more parsimonious interpretation. This is the reason for the three experiments reported here.

Experiment I

In Experiment I, the typical order of recall was reversed to investigate the relation between order and accuracy of recall. (The procedure differed from that used by Sampson and Spong (1961*b*) in that binocular fixation was not required.) It was also of interest to see if perceptually similar displays, projected monocularly and binocularly, resulted in stimuli being ordered similarly in recall, and to examine whether there were differences in accuracy of recall between the eyes when stimuli were viewed monocularly.

Method

Subjects. Ten students (five male and five female) from a first year psychology course were tested. Their ages ranged from 18–28 years (Mean = 21 years) and they were all right-handed and right eye dominant. Handedness was determined by an individually administered questionnaire composed of 20 items suggested by Humphrey (1951) as sensitive indices of manual preference, and three additional items designed by Gillies, MacSweeney and Zangwill (1960). Sighting preference was determined by the finger-aiming test (Buxton and Crosland, 1937). A Snellen chart was used to check visual acuity.

Apparatus. A magazine-load automatic projector was connected to two interval timers permitting control of stimulus duration and interval between successive stimuli.

Slides were "gravity-fed" into the projector from a vertically placed slide holder. The timers controlled the shutter mechanism of the projector. A standard tape recorder was used to record the subject's verbal responses. A binocular viewer, 106·7 cm. × 15·2 cm. × 7·6 cm., painted matt black, had a centre division down its length that effectively bisected the visual field. There was a day-view screen across the end of the viewer upon which stimuli for each eye were projected. The dimensions of this screen were 12·9 cm. × 6·3 cm., and the width of the centre division itself was 3 mm. The visual stimuli were arabic digits from 1 to 9, which had a projected size of 2·2 cm. × 1·3 cm. (subtending on the retina a horizontal angle of 41′ and a vertical angle of 1° 11′). They were mounted on 5·1 cm. × 5·1 cm. glass slides in pairs, in four different arrangements. These were:—

(a) *Left monocular* (*L.M.*): both digits mounted on the left-hand half of the slide, with one digit in the centre of the upper half of the slide and the other in the centre of the lower half of the slide.

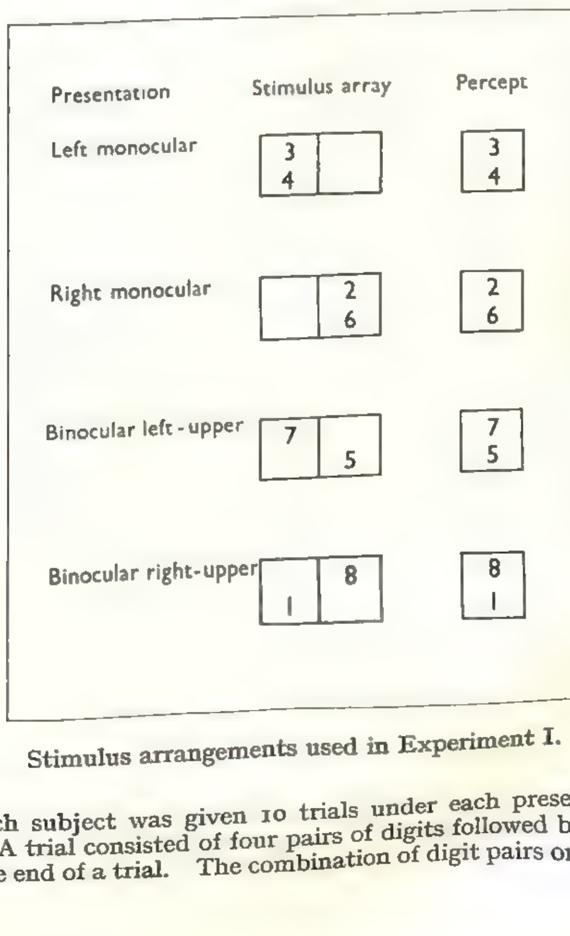
(b) *Right monocular* (*R.M.*): both digits mounted as in (a) but on the right half of the slide instead of the left.

(c) *Binocular left-upper* (*L.U.*): one digit mounted in the centre of the upper left-hand quarter of the slide and the other in the centre of the lower right-hand quarter of the slide.

(d) *Binocular right-upper* (*R.U.*): this was the reverse of condition (c), i.e. the right eye digit mounted in the centre of the upper right-hand quarter of the slide.

Twelve random pairs of digits were used in each presentation condition. The four different arrangements of the stimuli are illustrated in Figure 1, together with the perceived appearance of each digit pair.

FIGURE 1



Stimulus arrangements used in Experiment I.

Procedure. Each subject was given 10 trials under each presentation condition, a total of 40 trials. A trial consisted of four pairs of digits followed by a blue flash on the screen signalling the end of a trial. The combination of digit pairs on a trial was identical

for each subject and on no trial did the same digit appear twice. A random order for individual trials was determined in advance for each subject.

The subject was instructed to call out all the digits he could remember, in any order, as soon as possible after the signal to respond. This signal was delivered 0·8 sec. after the last stimulus pair on a trial. The interval between the end of each pair of stimuli and the onset of the next pair was also 0·8 sec., while the duration of each stimulus pair and the blue flash was 0·3 sec.

Results

A. *Percentage correctly recalled.* A response was scored correct if it corresponded to a stimulus, regardless of its position in recall. All accuracy data were transformed angle = arcsin $\sqrt{\text{per cent. correct}}$ (Snedecor, 1956) before analysis. In each presentation condition stimuli could appear in an upper or a lower position on the screen, and for purposes of analysis each position was considered to be a separate treatment. As there were four presentation conditions in all, this meant that there were actually eight stimulus conditions selected for comparison. Differences between the eight treatment sums were tested by a series of planned orthogonal comparisons (Hays, 1963). The particular comparisons of interest were: (1) Monocular vs. binocular performance; (2) Left monocular vs. right monocular; (3) Binocular left-upper; left vs. right eye (upper vs. lower position), and (4) Binocular right-upper; left vs. right eyes (lower vs. upper position).

TABLE I

PERCENTAGE OF TRIALS ON WHICH THE ORDER OF RECALL WAS AN UPPER-LOWER PAIRING OF RESPONSES, A LOWER-UPPER PAIRING, OR A MISCELLANEOUS ORDER WHERE NO SYSTEMATIC PAIRING WAS OBSERVED. (EXPERIMENT I)

Presentation condition	*U–L pairing*	*L–U pairing*	*Miscellaneous*
L.M.	68		32
R.M.	62	2	36
L.U.	71		29
R.U.	20	15	65

B. *Order of recall.* Table I shows digits were typically reported in pairs with the stimulus in the upper position on each slide recalled first in each pair (55·3 per cent. of all trials). The one exception to this general trend occurred when stimuli in the upper part of the screen were projected to the right eye (R.U. condition). On only 20 per cent. of trials under this condition was there upper-lower pairing in recall.

Spong (1961) made the observation that, if responses were organized temporally in pairs with the first response in each pair corresponding to a digit projected to the left eye, the numerical order of responses would be such that responses to digits projected to the left eye would be consistently "odd" numbered and responses to digits projected to the right eye would be consistently "even" numbered. Under the L.U. condition this clearly was the case. However, the majority of "odd" numbered responses with R.U. presentation corresponded to stimuli projected to the *right* eye. Thus this analysis shows that R.U. presentation was associated with

a right eye, left eye pairing of responses as the dominant order in recall, but that this tendency was far less marked than was the tendency towards left-right pairing in the L.U. presentation condition. Additional evidence for a different recall order under the R.U. condition is that 98 per cent. of all first responses in the L.U. condition corresponded to digits projected to the *left* eye, whereas in the R.U. condition 60 per cent. of all first responses corresponded to digits projected to the *right* eye.

Discussion. There were no differences between the eyes in accuracy of recall when viewing was with one eye alone, so there is no suggestion from this experiment that such differences can be demonstrated without binocular stimulation. Under monocular viewing more information must be handled from one eye, because the same number of digits was presented on each trial, whether monocular or binocular. Yet this additional information load had no effect, which indicates that simultaneous binocular stimulation has a significance other than that of simply adding to the number of stimuli which must be handled at once.

There are still difficulties in the interpretation of the role of bilateral stimulation however, since even in the two binocular viewing conditions no differences in accuracy of recall were associated with the eyes. This provides further evidence that the situations in which stimuli projected to the right eye are recalled more accurately than stimuli projected to the left eye are rather circumscribed. With this latter result, of course, it is impossible to state any conclusions about the relation between order and accuracy of recall.

One apparently anomalous finding was that while on 15 per cent. of all trials in the R.U. presentation condition the order of recall was lower-upper (left-right), there were no instances of a corresponding right-left order with L.U. presentation. A tentative explanation for this result is that the binocular viewing situation may provide a cue for a left-right as well as an upper-lower order of recall. A slight left-right displacement of the binocularly presented stimuli could constitute such a cue. Under the L.U. presentation condition left-right and upper-lower ordering tendencies would have the same effect, but in the R.U. condition both tendencies could not co-exist. The viewing condition is such that a displacement seems highly probable, for the distance between the mid-points of each monocular screen (6·6 cm.) is slightly greater than the average inter-pupillary distance. This means that the necessity to diverge the eyes in order to fuse the separate monocular views would run counter to the tendency to make use of accommodative convergence (Borish, 1954) with the result that the eyes would be slightly "over-converged" in relation to the stimuli being viewed, a stimulus projected to the left eye appearing slightly to the left of its true position and a stimulus projected to the right eye appearing slightly to the right of its true position.

Experiment II

The situations in which there has been more accurate recall of information projected to the right eye, have been those in which stimuli were projected to the centre of *each* monocularly viewed screen. Because of the fusion of the two monocular views these stimuli are seen as overlapped, a feature that was absent in the binocular viewing conditions of Experiment I. Experiment II replicates this overlap condition, and also simulates the binocular percept monocularly to test for differences in accuracy of recall when overlapped stimuli are viewed monocularly.

Method

Subjects. Ten students (five male and five female) were selected according to the criteria employed for Experiment I. Their ages ranged from 18–22 years (Mean = 19·4 years).

Apparatus. The same apparatus was used as in Experiment I except that the visual stimuli were rearranged so as to create the following presentation conditions:—

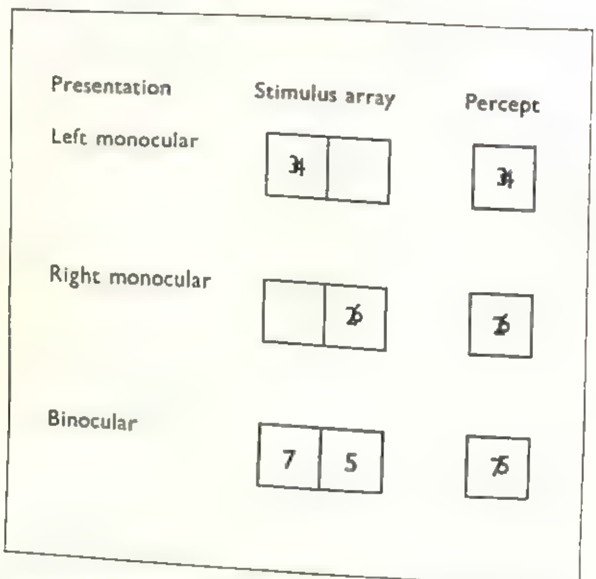
(*a*) *Left monocular* (*L.M.*)*:* both digits mounted on the left-hand half of the slide with a 50 per cent. overlap. The overlapped pair was positioned so that it appeared in the centre of the left monocular field.

(*b*) *Right monocular* (*R.M.*)*:* both digits mounted as in (*a*) but on the right-hand half of the slide.

(*c*) *Binocular* (*Bn.*)*:* the left and right eye digits mounted in the centre of each half of the slide.

As the binocular presentation condition encompassed both the L.U. and R.U. conditions of Experiment I there were 24 random pairs of digits to be used in this condition and 12 pairs for each of the monocular conditions. There were six stimulus conditions in all. The three different arrangements of the stimuli are illustrated in Figure 2.

FIGURE 2



Stimulus arrangements used in Experiment II.

Procedure. The procedure followed that of Experiment I. The number of trials under binocular viewing was not reduced from 20 because there was only one binocular presentation condition. This ensured that the ratio of monocular to binocular trials was the same as in the earlier experiment and the same number of stimuli were delivered to each eye under monocular and binocular viewing.

Results

A. Percentage correctly recalled. The results were analysed as in Experiment I. The comparisons selected for evaluation in the accuracy analysis were: (1) Monocular vs. binocular performance; (2) Left monocular vs. right monocular; (3) Binocular; left vs. right; (4) Left monocular; left vs. right stimulus position, and (5) Right monocular; left vs. right stimulus position.

In accordance with the findings of previous studies (Sampson and Spong, 1961*a*, 1961*b*), there was significantly (F, 1, 45 = 4·35, $p < 0·05$) more accurate recall in the binocular viewing condition of digits projected to the right eye. Recall was also significantly more accurate (F, 1, 45 = 6·64, $p < 0·05$) when viewing was monocular

than when viewing was binocular. There were no significant differences in accuracy of recall between the eyes when viewing was monocular.

B. Order of recall. Stimuli were generally reported in the order presented, with the left-hand stimulus in each pair given first. The tendency to left-right pairing was *more* marked in the monocular presentation conditions.

In the L.M. condition 90 per cent. of all first responses corresponded to a stimulus appearing in the left-hand position; the figures for the R.M. and Bn. conditions are 87 per cent. and 61 per cent. respectively.

Discussion. The results of the statistical analysis, together with the spontaneous subjective reports of the subjects, indicate that the desired overlap of stimuli was obtained in the binocular viewing condition, and it is apparent from the statistical analysis that the difference between the eyes in binocular performance is related to a competition between the inputs from each eye, which is resolved in favour of right eye information. Support for this conclusion is seen in the fact that the overall difference between levels of monocular and binocular performance can be attributed almost entirely to the reduced accuracy of recall, under the binocular presentation condition, of information projected to the *left* eye. The accuracy with which information projected to the right eye was recalled was practically identical under both monocular and binocular viewing. Since no eye difference under monocular viewing was observed, the binocular eye differences cannot be attributed to an inherent weakness of the left eye itself as Sampson and Spong (1961*b*) pointed out earlier.

It might be argued however that similar results could be obtained if the subjects had heterophoria accompanied by low fusional vergences, in which case they might have seen the letters projected binocularly in positions somewhat different from that illustrated in Figure 2. However, this would be expected to affect *order* of recall rather than accuracy of recall since all subjects had normal visual acuity. Further, since the walls of the viewing tube served as very strong stimuli to fusion, only subjects with extreme binocular visual anomalies (which was not the case here) would be expected to have the eyes directed in other than a slightly "over converged" position.

The fact that there were no differences in accuracy between the eyes under monocular conditions shows that stimulus overlap without binocular involvement is not sufficient to reveal any such differences. Previous studies (Teuber and Weinstein, 1956; Milner, 1958; Sinha, 1959; Kimura, 1963) have demonstrated the sensitivity of performance of subjects with cerebral lesions to conditions in which the test material has been presented in a context which makes for perceptual difficulty. The "perceptual" overlap with *binocular* viewing here, may have a similar effect, but the absence of any eye differences with *monocular* overlap emphasizes the importance of the binocular component—a complex stimulus display on its own is not sufficient to bring out any eye differences.

Order of recall under the two monocular presentation conditions reinforces the suggestion from Experiment I that a slight lateral displacement of the stimuli is sufficient to sustain a left-right order in recall. Another finding of importance is that stimuli viewed monocularly were recalled as efficiently when they appeared displaced towards the left as when the displacement was in a right-hand direction. Thus there is no evidence from this experiment that order and accuracy of recall are related. This is emphasized by the fact that left-right recall order was far less stable in the binocular viewing condition, which was the only situation in which there was more accurate recall of the right-hand stimulus.

Experiment III

The third experiment seeks differences in accuracy between the eyes in a context other than this memory task. The findings of Experiment II suggest that with stimulus overlap, rivalry between information originating at each retina may be such that one stimulus is seldom seen. One classic feature of binocular rivalry is that the whole figure presented to one eye may be suppressed, not merely the part that corresponds to points stimulated on the other retina (Helmholtz, 1962). Experiment III examines the possibility that a process of this nature may be associated with differences between the eyes in accuracy of recall.

Method

Subjects. Ten students (five male and five female) were selected according to the criteria employed for Experiments I and II. Their ages ranged from 17–22 years (Mean = 19·1 years).

Apparatus. The apparatus was unaltered for this experiment. The visual stimuli were the 24 random pairs of digits used on the binocular trials in Experiment II.

Procedure. The 24 pairs of digits were presented once in random order to each subject. They were instructed to call out what they saw as each slide was presented. The interval between the stimulus pairs was 0·8 sec., and each pair was projected on the screen for 0·3 sec.

Results and Discussion. Percentage correct for left and right eyes was calculated for each subject and the scores transformed Angle = Arcsin $\sqrt{\text{per cent. correct}}$ before analysis. The mean level of accuracy of report was 74·2 per cent. from the left eye and 85·0 per cent. from the right eye. A one-tailed t test for correlated means shows that this difference is statistically significant ($t = 1{\cdot}94$, $d.f. = 9$, $p < 0{\cdot}05$). It will be recalled that it was shown in Experiment II that the one condition required to demonstrate accuracy differences between the eyes was an unstable binocular percept. Experiment III suggests that under these conditions there is binocular rivalry, regarded here as a cortical response to one or other of the competing fields (Woodworth and Schlosberg, 1954; Bower and Haley, 1964), which is resolved in favour of the right eye. On the other hand when subjects are required to *fixate* binocularly (Sampson and Spong, 1961*b*) superimposing two points while digits are projected, there are no differences in accuracy of recall between the eyes.

Taken altogether, these results suggest that providing fixation points binocularly results in an equal division of attention between the two eyes in subsequently presented material, which does not occur when the two eyes are left free. The results of the present experiments, in view of the handedness and eye dominance of the subjects, are also consistent with the hypothesis that the outcome of the binocular rivalry in Experiment III resulted from these subjects being left dominant cerebrally.

This research was supported by an equipment grant from the University Grants Committee.

References

Borish, I. M. (1954). *Clinical Refraction.* Chicago: Professional Press.

Bower, T. G. R., and Haley, L. J. (1964). Temporal effects of binocular vision. *Psychon. Sci.*, **1**, 408–10.

Buxton, C. E., and Crosland, H. R. (1937). The concept of "eye preference." *Amer. J. Psychol.*, **49**, 458–61.

Gillies, S. M., MacSweeney, S. A., and Zangwill, O. L. (1960). A note on some unusual handedness patterns. *Quart. J. exp. Psychol.*, **12**, 113–6.

Hays, W. L. (1963). *Statistics for Psychologists.* New York: Holt, Rinehart & Winston.

HELMHOLTZ, H. VON (1962). *Treatise on Physiological Optics* (Vol. III). SOUTHALL, J. P. (Ed.). New York: Dover.

HUMPHREY, M. E. (1951). Consistency of hand usage. *Brit. J. educ. Psychol.*, **21**, 214–25.

KIMURA, D. (1963). Right temporal lobe damage: Perception of unfamiliar stimuli after damage. *Arch. Neurol.*, **8**, 264–71.

MILNER, B. (1958). Psychological defects produced by temporal lobe excision. *Publ. Proc. Res. Nerv. Ment. Dis.*, **36**, 244–57.

SAMPSON, H. (1964). Immediate memory and simultaneous visual stimulation. *Quart. J. exp. Psychol.*, **16**, 1–10.

SAMPSON, H., and SPONG, P. (1961*a*). Handedness, eye-dominance and immediate memory. *Quart. J. exp. Psychol.*, **13**, 173–80.

SAMPSON, H., and SPONG, P. (1961*b*). Binocular fixation and immediate memory. *Brit. J. Psychol.*, **52**, 239–48.

SINHA, S. P. (1959). *The Role of the Temporal Lobe in Hearing*. Unpubl. Masters thesis, McGill University.

SNEDECOR, G. W. (1956). *Statistical Methods.* Iowa: Iowa State College Press.

SPONG, P. (1961). *Cerebral Dominance and Immediate Memory.* Unpublished M.A. thesis, Univ. of New Zealand.

TEUBER, H.-L., and WEINSTEIN, S. (1956). Ability to discover hidden figures after cerebral lesions. *Arch. Neurol. Psychiat.*, **76**, 369–79.

WOODWORTH, R. S., and SCHLOSBERG, H. (1954). *Experimental Psychology.* New York: Holt.

Manuscript received 3rd March, 1967.

PERIMETRIC STUDY OF FIELD DEFECTS IN MONKEYS AFTER CORTICAL AND RETINAL ABLATIONS

BY

ALAN COWEY

From the Psychological Laboratory, University of Cambridge

Monocular visual field defects were studied in two monkeys. In one, the macular retina was destroyed by photocoagulation, producing a central scotoma and consistent 5° eccentric fixation. In a second animal the effects of removal of macular projection area in striate cortex and subsequent photocoagulation of the macula were compared. The cortical operation produced a partial field defect, i.e. a region of diminished sensitivity but not a scotoma, which became with practice much smaller than the region of retina whose primary projection area had been ablated. A 10° eccentric fixation was observed. Following the second, retinal, operation a macular scotoma was demonstrated whose size and position corresponded closely with the area of retinal destruction as determined by photography of the fundus and later histological examination of the retina.

Introduction

In a previous perimetric study of visual field defects in monkeys following removal of various portions of striate cortex it was reported that although the defects were roughly of the expected size, shape and position, they were not regions of blindness (Cowey and Weiskrantz, 1963). In each case the ablation produced an amblyopia not a scotoma. A possible objection to this earlier conclusion is that the animals may have been detecting the stimuli (brief flashes of light subtending 1°) by scattered light. The authors did not control for this artefact by plotting the natural blind spot and showing that it was indeed blind. This essential control is performed in the present study.

It was further found in the earlier study that each animal's ability to detect the stimulus in the defective region gradually improved, until a stimulus intensity which initially revealed a defect was eventually totally ineffective for this purpose. This improvement was observed in the first few months after operation. It was not possible to decide whether the improvement was spontaneous or the result of practice. In the present study the post-operative testing with one animal did not begin until $2\frac{1}{2}$ years after removal of lateral striate cortex, by which time any spontaneous recovery was presumed to be complete.

There was some indication in the previous study that macular field defects were smaller than would be predicted from the extent of ablated striate cortex, the latter being compared with electrophysiological maps of the retinal projection to cortex. Unfortunately the defects were plotted with fixed stimuli set at 5° intervals in a perimeter, and the precise determination of the boundaries of a defect was impossible. In the present study the defects are plotted with movable as well as with fixed stimuli, and an attempt is made to study changes in size as well as in sensitivity of an impaired region.

Perimetric studies of field defects caused by retinal damage have not previously been reported in monkeys. They are of special interest here for two reasons. First, a region of retina in which the receptors are totally destroyed should be blind, providing an excellent control for possible artefactual effects of scattered light without having to rely on the natural blind spot which, being small, is difficult to plot in an

animal whose fixation cannot be controlled or measured with the accuracy common in human clinical perimetry. Second, the position and angular extent of retinal destruction can be accurately determined by photography of the fundus of the living animal and subsequent histological examination; therefore the behavioural effects of destroying a known region of retina can be compared with removal of striate cortex corresponding, according to electrophysiological and anatomical maps, to a known region of retina. This comparison is studied here, in one case by performing the two types of operation sequentially on a single animal.

Finally, evidence of eccentric fixation following striate cortex damage was reported in the earlier paper. A more successful method of attracting and retaining a monkey's visual attention is described here, and is used to demonstrate eccentric fixation following striate cortex or retinal injury. The fact that the natural blind spot can be plotted is also used to reveal eccentricity of gaze, for if any point other than the fovea is used for fixation the natural blind spot appears to be displaced from its normal position with respect to the fixation mark.

Method

Each animal was taught to press its eye to a peep-hole, look at a perimeter, and await the presentation by the experimenter of either of two kinds of stimulus. The first consisted of a 50 millisec. click from a buzzer, and the animal was automatically rewarded with a food pellet if he pressed a lever to his right. If he pressed the lever to his left the light above the testing cage was automatically extinguished for 5 sec. The other stimulus consisted of the same click and a coincident flash of light from any one of 121 stimulus bulbs set in the perimeter. (See Cowey and Weiskrantz (1963), for details of stimulus bulbs.) The animal was rewarded for pressing the left lever in response to this paired stimulus and punished as before if he pressed the lever to his right. The paradigm is: click—go right, click plus light—go left. The animal's viewing eye was photographed on every trial and his direction of gaze on any trial could be subsequently determined by comparing the photograph, which included a trial counter, with a set of standard reference photographs consisting of a record of each animal fixating up to 64 points in the perimeter, spaced at 5° intervals (see Cowey, 1963; Cowey and Weiskrantz, 1963; for further details of this method). Field defects were plotted by analysing the eye photographs of trials on which the animal responded incorrectly to the paired stimulus, and relating the position of the undetected flash to the direction of gaze.

Apparatus

A full description and diagram of the apparatus appears elsewhere (Cowey and Weiskrantz, 1963). Briefly, it consisted of a curved screen, part of a sphere of radius 50 cm., in which 1° stimulus bulbs were set at 5° intervals. The entire screen covered about 80° of visual field. At the edges of the screen were four 60 watt floodlights. They provided general illumination for photography, corneal reflexions for determining locus of fixation, and they ensured that the animal was light-adapted. A peep-hole, at the centre of curvature of the perimeter, allowed the animal to view the display with his right eye, and all visual field defects reported in this paper refer to the right eye only.

Several changes were made in the stimulus arrangements described in the earlier paper (Cowey and Weiskrantz, 1963). First, the luminance of the stimuli was varied through 3 log units, from 5,500 to 5·5 cd./m^2, by placing individual neutral density filters over the bulbs. These luminance values, which refer to peak intensity of the flash, are slightly different from those described before, because of a voltage change in the control apparatus. Luminance of the flood lights was unchanged at 17,000 cd./m^2. Second, the angle subtended by a stimulus could be reduced from 1° to ½° by masks placed over the bulbs. Third, once field defects had been roughly plotted, their boundaries were accurately determined by using one or more bulbs which could be moved in 1° steps across the perimeter.

Training and testing

The animals required from 6 to 9 months of daily testing, with up to 200 trials per session, before they reliably performed with better than 95 per cent. correct responses.

A trial was started by the experimenter, who operated the camera switch when the animal was looking through the peep-hole. The two kinds of trial were presented in random order by means of two 4-bank 25-position uniselectors which together provided 400 independent outputs. Since a different starting point in the sequence was chosen each day and the entire sequence was changed every few weeks, neither the animal nor the experimenter could predict whether the next trial would include a visual stimulus. Nor could the position of a visual stimulus be predicted.

When the movable bulbs were used to plot a field defect it was arranged that the stimulus should not fall within the defect on more than about one trial in 20. It is possible to do this of course only if an animal's fixation can be controlled and after a defect has been roughly determined. Such infrequent presentation of crucial stimuli makes plotting a defect a tedious and lengthy procedure but is essential, for the performance of an animal which is forced to make errors is likely to deteriorate. For this reason it took several weeks or months to plot even a small defect.

Standard reference photographs and control of fixation

These were obtained by mounting a small plane mirror, subtending 2°, at various positions in the perimeter. It was mounted in such a way that the animal would see the reflexion of his own eye, and proved to be an excellent device for attracting and retaining an animal's attention. The mirror was therefore retained in the centre of the perimeter for all testing reported here and was so successful in controlling fixation over long periods of time that the majority of eye photographs were identical, and many of the standard reference photographs were never used. The only exception to this central positioning of the mirror occurred during post-operative tests of eccentric fixation when the mirror was again placed at various positions in the perimeter and the procedure for obtaining standard reference photographs was repeated.

One further strategem was used to control fixation. Since the experimenter observed the animal's eye before every trial began it was possible to withhold the stimulus if the animal was clearly not fixating the central mirror, e.g. if he was looking more than 10° away in any direction. It is very likely therefore that the animal learned to look at the mirror in order to start a trial.

Subjects and surgery

Two rhesus monkeys, one male and one female, were used. They had previously sustained bilateral removal of inferotemporal cortex without effect on perimetric performance (Cowey and Weiskrantz, 1967). They were still immature at the beginning of the present experiment but had matured and weighed about 8 kilos by the time it ended.

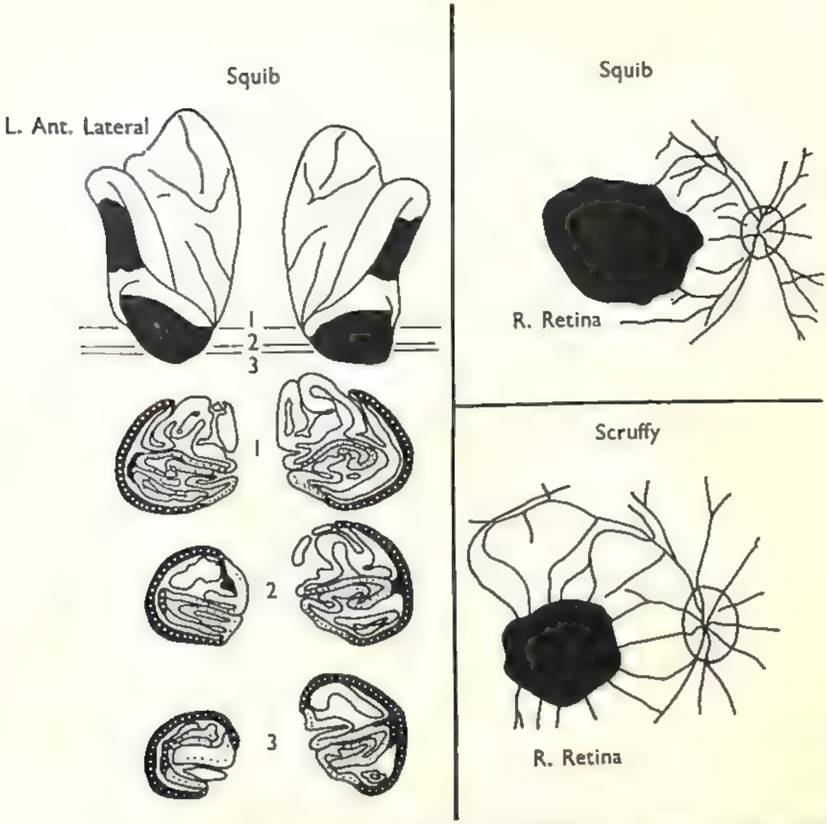
In one monkey (Squib) the striate cortex corresponding to the first 8°–10° of central retina (Talbot and Marshall, 1941; Daniel and Whitteridge, 1961) was removed bilaterally by sub-pial suction. Three and one half years later the macular region of the retina (7°–8° radius) was destroyed in both eyes, using a Zeiss Xenon-arc photocoagulator. The latter operation was also performed on the second animal (Scruffy). All operations were carried out under deep anaesthesia (Nembutal: sodium pentobarbitone).

Histology

At the end of the experiment the animals were anaesthetized and the fundus of each eye was photographed from several angles with a Zeiss retinal camera. The animals were then killed with an injection of Nembutal, and perfused through the heart with normal saline followed by formal-saline. Paraffin sections were cut at 20 μ of the brain of one animal (Squib), and every 20th section stained with thionin. The eyes of both animals were embedded in paraffin and sectioned at 6 μ before being stained with haematoxylin-Ponceau acid fuchsin according to Masson's method.

Histological reconstructions are shown in Figure 1. The inferotemporal damage is not relevant to this paper and is discussed elsewhere (Cowey and Weiskrantz, 1967). Degeneration in the lateral geniculate bodies is not shown and will be discussed in a separate study of the effects of retinal lesions on acuity. The destruction of lateral striate cortex in Squib is complete and extends somewhat on to the medial surface. There is also unintended damage to striate cortex in the underlying calcarine fissure. This is shown in the drawings of coronal sections and is important in understanding the field defects to be described. The retinal lesions for the right eye of each animal are traced from colour transparencies of the fundus. The edge of the lesion was very sharp in the photographs and sections through the lesions showed that the receptors were totally

FIGURE 1



Reconstructions of brain lesions and retinal destruction. Areas of destruction are shown in solid black, with white stippling to indicate absent striate cortex on the representative cross sections of the brain. Intact striate cortex is shown by black stippling.

destroyed within the area shown in black. Further details about reconstructing the exact position of the lesion with respect to the fovea are given by Weiskrantz and Cowey (1967) in a study of the effects of retinal damage on visual acuity.

Results

Pre-operative testing

By pre-operative is meant testing before and after the bilateral removal of inferotemporal cortex, which had no effect on visual fields as tested with stimuli varying from 5,500 to 5·5 cd./m². Over several thousand trials both animals scored better than 90 per cent. correct with the dimmest stimulus, rising to 95–97 per cent. with the brightest. False negative errors, i.e. failing to report a visual stimulus, were slightly more common than false positives, i.e. reporting a non-existent visual stimulus. The distribution of false negatives was not uniform over the visual field: there were more above the visual axis than below it, and more to the right than to the left of it (excluding the special case of the natural blind spot which of course lies to the right). These observations simply confirm what has been previously reported in normal animals (Cowey and Weiskrantz, 1963).

No attempt was made pre-operatively to plot the natural blind spot in one animal (Squib), but in the other (Scruffy) the centre of the blind spot was detected

quite easily with a 55 cd./m², $\frac{1}{2}$° stimulus moved in 1° steps across the relevant part of the field. It lay $16\frac{1}{2}$° to the right of the visual axis. When the stimulus was presented in this position the animal made 87 per cent. false negative responses. Ten degrees above or below this point there were only 3 per cent. false negatives. When the stimulus was moved 1° in any direction from the centre of the blind spot the number of false negatives fell to about 50 per cent., 3° away they were barely above the control level for normal parts of the field. The blind spot is known from anatomical measurements to be about 7° by 5° (and these measurements were confirmed in two other anaesthetized animals with a measuring ophthalmoscope). The results therefore indicate that, at least in this animal and with the given stimulus, an absolute blind area is 2°–3° larger than the region whose borders are defined by 50 per cent. false negatives, and perhaps 4° larger than the region where there are 87 per cent. false negatives. This result is important in evaluating the size of the field defects reported below. It presumably occurs because of scattered light, which is detectable at the edge of a defect, and because in a method where fixation is given to the nearest 5° point there is always the possibility of a $2\frac{1}{2}$° error.

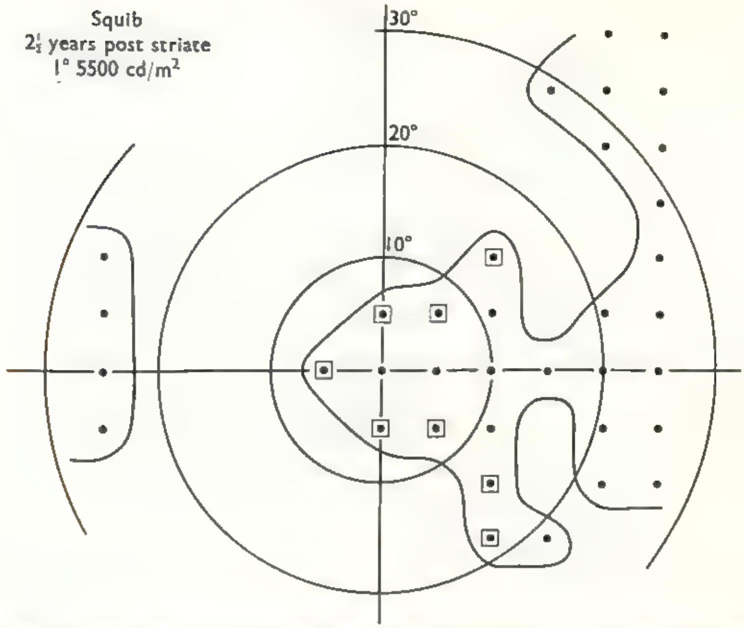
In all subsequent Figures where the natural blind spot is plotted it is shown on the horizontal retinal meridian. This is not quite correct, for of course the natural blind spot should lie slightly below it. However, since there was no means of determining exactly where the horizontal retinal meridian was represented in the visual field it has been drawn through the plotted blind spot.

Post-operative testing

(1) *Effects of lateral striate cortex removal on visual field* (*Animal Squib*). The operation was performed in March, 1961. Testing was resumed in October, 1963, allowing a comparison with the immediate effects of this operation on perimetric performance (Cowey and Weiskrantz, 1963). The results of the first 14,000 trials, using the brightest 1° stimulus are shown in Figure 2. Defective points are shown by black dots; the squares enclosing certain dots are explained later and may be ignored in the first instance. The defect was plotted with the entire perimetric array of fixed stimuli. To the left of mid-vertical the central defect extends about $7\frac{1}{2}$° along all meridia, which is entirely consistent with the electrophysiological map of the retinal projection to lateral striate cortex. The island defect about 25° to the left was unexpected, but is explained by the accidental damage to striate cortex in the underlying calcarine fissure (see Fig. 1). The position of this defect is also consistent with the admittedly more approximate map of the projection of peripheral retina on to calcarine cortex, which shows that the horizontal retinal meridian is represented in the lateral fold of the calcarine fissure (Daniel and Whitteridge, 1961). To the right of the vertical mid-line the field defect can be considered in two parts. The more central (macular) portion is larger than its counterpart on the left and this is consistent with the fact that the lesion in the contralateral left hemisphere extended on to the medial surface of the brain. The extensive and peripheral defect on the right was also unexpected and again correlated with damage to the calcarine fissure. The furthest borders of the two peripheral defects were not determined, and it is not certain that they were insular. Insular field defects would be predicted from the histology. The apparent bridge between macular and peripheral defects on the right is the natural blind spot, as revealed by more accurate plotting, and as shown in subsequent Figures.

The severity of the impairment is not uniform throughout the defective regions Outside the defect the ratio of detected to undetected flashes was 97 : 1, within the

FIGURE 2



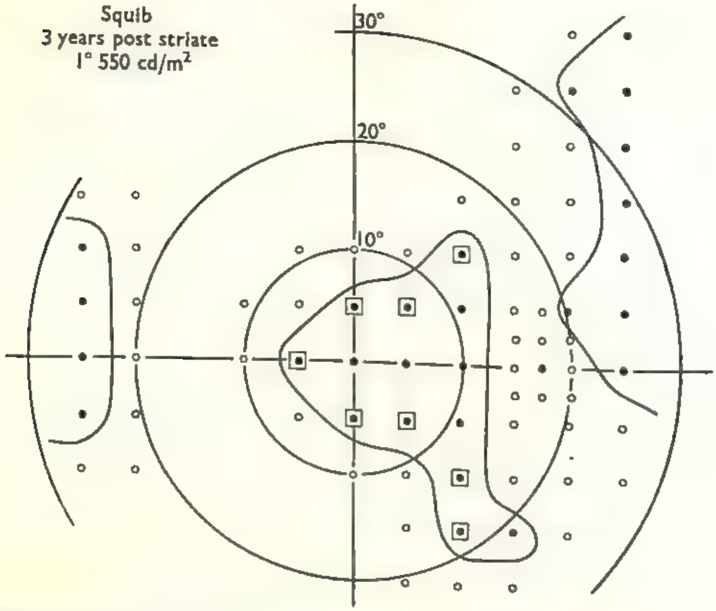
Field defects plotted with a 1° stimulus of maximum intensity approximately 2½ years after bilateral removal of lateral striate cortex. Defective points are indicated by black dots. Elsewhere, over a region 50° across, responses were normal. Data taken from 14,000 trials over a 4-month period during which each stimulus of an array of 100 was presented 70 times. Luminance values refer to peak intensity of a 50 millisec. flash.

defects as a whole it was 9 : 1. In the centre of the macular defect 5° to the right of the visual axis the ratio was 5·5 : 1 whereas at the edge of the defect, at points marked ⊡, it was 11 : 1. The greatest impairment existed in the large peripheral defect where the ratio was 3 : 1, 25° to the right of the visual axis and 5° above it. This point is close to one of the floodlights used to illuminate the eye, which would raise the ambient illumination in the corresponding region of the retina.

The percentage of false negative responses declined throughout the period of testing until the defect was barely detectable, *except* at the natural blind spot. This recovery parallels that already reported in animals soon after operation (Cowey and Weiskrantz, 1963).

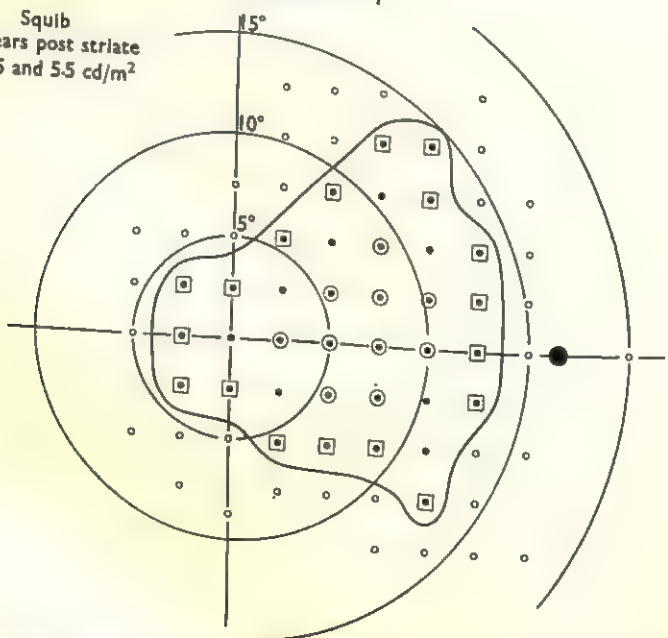
At this point the stimuli were dimmed by 1 log unit and the defect plotted again. The results are shown in Figure 3. Note that the defect has retreated markedly at bottom and top right of the temporal insular defect, as compared with Figure 2. Outside the defect 1 in 35 flashes were misreported. Just inside, or at, the border of the macular defect, at points marked ⊡, the ratio of detected to undetected flashes was 5·5 : 1 whereas in the centre of the defect, 5° from the visual axis, the ratio was 2 : 1. In the periphery, 25° to the right of the visual axis and 5° above the ratio was 1 : 1, i.e. the stimulus was at threshold intensity *at this time*. The region of the optic disc was explored with an array of nine bulbs set 2½° apart in a 9 × 9 array. At only one of the positions, which is the approximate position of the centre of the natural blind spot, was there strong evidence of a defect. This natural defect, despite being small, is more severe than any part of the induced defect. The ratio of detected to undetected stimuli at this position was 0·85 : 1. Once again the number of false negative responses declined during testing, except at the natural blind spot, until it was very difficult to demonstrate any impairment.

FIGURE 3



Field defects plotted approximately 3 years after striate cortex removal and following reduction in stimulus intensity from 5,500 to 550 cd./m². Defective points indicated by black dots. Positions close to the defect and at which responses were normal are indicated by circles. Data from 7,000 trials over a 2-month period. For further description see text.

FIGURE 4



Central field defect plotted over a 2-month period approximately 3½ years after lateral striate cortex ablation, and using a movable stimulus, further reduced in size and intensity. Symbols as in previous figures. The large solid black circle indicates the centre of the natural blind spot. Note change in scale as compared with preceding two Figures. For further details see text.

At this point further changes were made in stimulus presentation. The macular region was investigated with a smaller, ½°, stimulus, further reduced in intensity by 1 and then 2 log units. This stimulus was moved across the macular region in 2½° steps, with fixation controlled as before by the small mirror. To ensure that performance did not break down because of the inevitable errors when this stimulus appeared in the defective region the vast majority of visual stimuli were kept at maximum size and luminance and presented in what earlier results had shown to be intact parts of the visual field. Performance with these stimuli was almost perfect. The results with dim stimuli are shown in Figure 4. At 55 cd./m² there were 142 correct and 77 incorrect responses, or 1·8 : 1, to a stimulus in the defective region as a whole. Just outside the defect, at points marked by circles, the figures were 117 : 16 or 7·3 : 1. At 5·5 cd./m² the number of detected and undetected stimuli within the defect was 77 : 119 or 0·65 : 1. Outside it was 130 : 7 or 18·6 : 1. This superiority in detecting the *dimmer* stimulus just outside the defect may seem anomalous until it is remembered that the animal constantly improved with testing and the results with the dimmer stimulus were the last to be obtained.

Once again the impairment is greatest at the centre of the defect. At points marked ⊙ the ratio of detected to undetected stimuli (combining results at both luminances) was 0·76 : 1, whereas at points indicated ⊡ the ratio was 1·25 : 1. Despite the faintness of the stimulus the macular defect has contracted as compared with Figures 2 and 3.

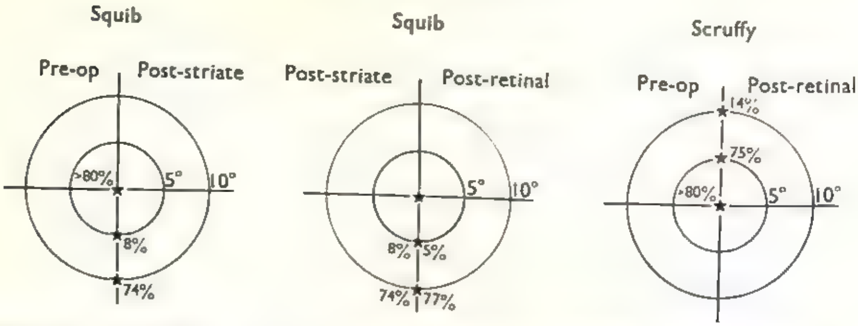
The final point of importance concerns the natural blind spot. This was studied at the same time with a ½° 55 cd./m² stimulus, maintained in such a position, relative to the fixation mirror, that it illuminated the centre of the optic disc. There were seven detected to 45 undetected stimuli, ratio 0·15 : 1. When corrected for false positive responses this came to 0·1 : 1, i.e. there were far fewer detections at the natural blind spot than in the field defect, indicating that the ability to detect a stimulus there is not an artefact of scattered light. No attempt was made to plot the total extent of the natural blind spot.

One result of the entire post-operative testing schedule is not shown in Figures 2–4. It was noted that both the macular field defect and natural blind spot were displaced downwards by about 10° from the fixation mirror, i.e. the animal was not using his fovea to fixate the mirror. This eccentric fixation, which was consistent throughout the entire post-operative testing period of 13 months was specifically studied at the end of this period. Results are described below.

(2) *Changes in fixation.* When the standard reference photographs were obtained before operation, more than 80 per cent. of the photographs for each of 64 positions of the fixation mirror were identical and were taken to represent foveal fixation. The remaining photographs indicated fixations scattered at random in the visual field. The procedure for obtaining the photographs was repeated 3½ years after operation but with fewer mirror positions. The result is shown in Figure 5, left. Post-operatively the animal "looked" with his fovea at a position 10° below the mirror on 74 per cent. of all trials, and directed his fovea at the mirror itself no more often than other positions scattered randomly in the field with respect to the mirror. This result confirms the impression of eccentric fixation revealed during perimetric testing.

(3) *Effects of retinal destruction following striate cortex removal* (*Animal Squib*). (*a*) *Field defects.* The macular region of the retina in both eyes was destroyed 3½ years after the cortical operation, permitting a study of the effects of the separate operations in a single animal. Perimetric testing was resumed 3 months later

FIGURE 5

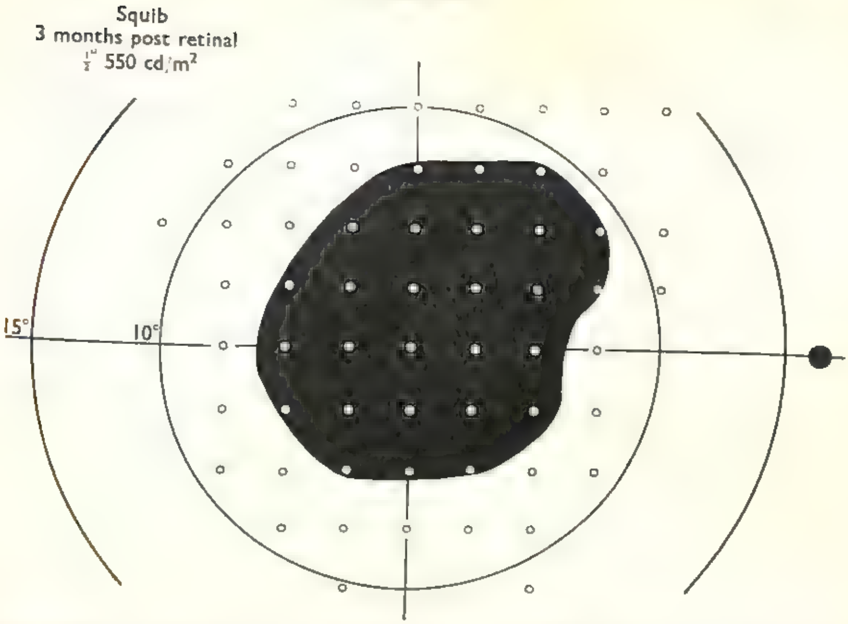


Distribution of fixations with respect to a small plane mirror placed at various positions in the perimeter. Pre-operatively (values to left of vertical meridian) both animals fixated the mirror on more than 80 per cent. of all occasions. After removal of lateral striate cortex, fixation was predominantly 10° below the mirror (diagram on left, data from 1,134 trials with eye stationary and using 15 different mirror positions). This pattern was maintained after subsequent macular destruction (centre diagram, data from 282 trials with eye stationary, and using four mirror positions). After destruction of macular retina in the second animal fixation was consistently elevated by 5° or more (diagram on right, data from 332 trials with eye stationary, using eight mirror positions). With both animals 10 to 20 per cent. of fixations were distributed randomly in the perimeter with respect to the mirror. These are not plotted.

according to the following procedure, which was used in all subsequent testing with both animals. The macular defect was plotted with a ½° stimulus moved in 2½° steps across the central visual field, fixation being controlled as before. As in earlier testing the majority of visual stimuli were presented at maximum luminance in intact portions of the field. The first results, obtained with a ½° 550 cd./m^2 stimulus, are shown in Figure 6. At points indicated by small black circles performance was better than 90 per cent. correct, despite the fact that some of these positions lie within the earlier defect produced by striate cortex damage. Solid white circles indicate points within the defect. The edge of the defect is defined by approximately 50 per cent. correct performance and here small undetectable errors in determining fixation will sometimes allow the stimulus to fall on intact retina. Inside the defect, i.e. at points away from the edge, there were nine detected stimuli compared with 93 failures and there was no improvement in performance throughout testing. About five "correct" detections would be expected by chance when the animal's continuing tendency to make false positive responses is taken into account. The effect of the operation is therefore to produce a true scotoma, i.e. a region of absolute blindness, as compared with the earlier cortical operation which produced only a partial defect or amblyopia. In the very centre of the scotoma the response to a 1° stimulus of maximum intensity was also studied. There were four detections compared with 47 errors. The scotoma was therefore detectable using a stimulus which 3 years after the cortical operation had completely ceased to reveal any defect.

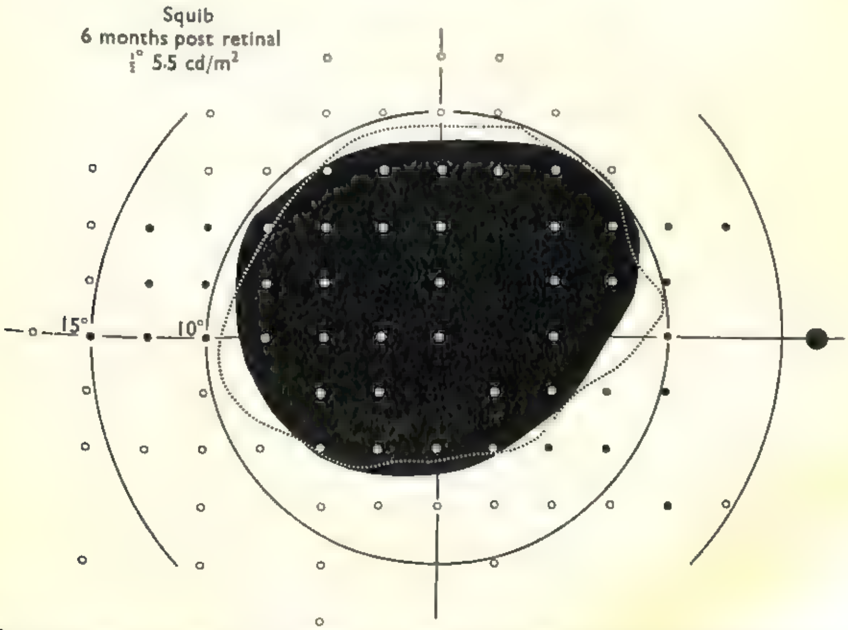
Six months after the retinal operation the effects of further reducing stimulus intensity were studied. Results are shown in Figure 7. With this stimulus the original defect caused by striate cortex damage was again detectable, hence the errors (solid black circles) immediately to the right of the scotoma. However, the defect remained partial at these points. At points entirely within the scotoma there were only eight detections compared with 106 errors. This ratio is very close to that for the centre of the natural blind spot (3:26), which was studied concurrently and with an identical stimulus. The following points about the scotoma are

FIGURE 6



Macular scotoma plotted over a 2-month period following destruction of central area of right eye. Scotoma is shown in black. Intact points indicated by circles, blind points by white dots. The centre of the natural blind spot was plotted concurrently using a ½° 55 cd./m² stimulus, as in Figure 4. For further details see text.

FIGURE 7



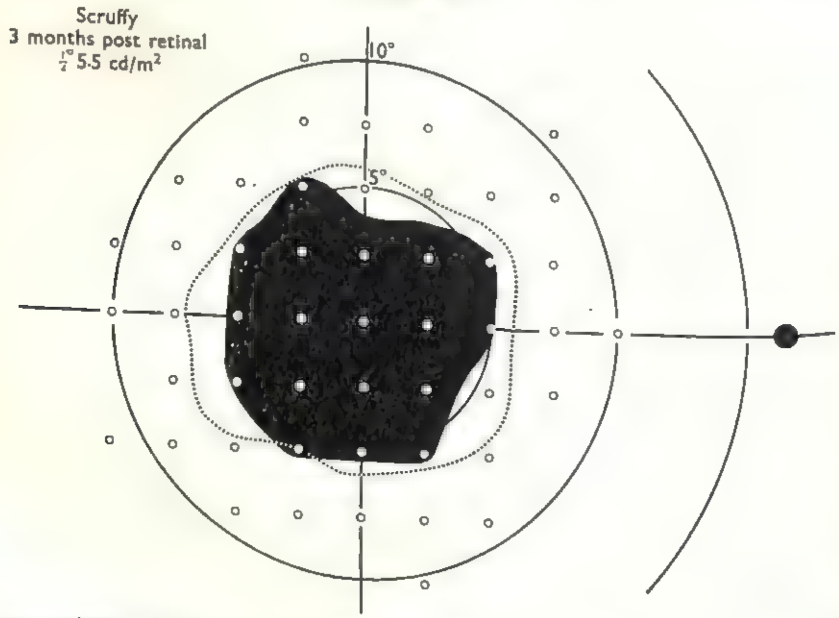
Macular scotoma plotted over a 2-month period approximately 6 months after operation and following reduction in stimulus intensity by 2 log units. The black dots indicate regions of partial defect outside the scotoma. Other symbols as in Figure 6. The dotted line superimposed on the scotoma indicates boundary of defect as predicted from retinal photography. For further details see text.

important: (i) The defect did not shrink with practice, as contrasted with the earlier amblyopia. It is slightly larger than in Figure 6, as would be expected with a dimmer stimulus which provides less scattered light. (ii) The shape of the defect closely fits the outline of the retinal lesion (dotted line) as determined by photography of the fundus (see Fig. 1). The photograph was inverted to correspond with retinal image inversion, and appropriately enlarged so that the distance fovea to centre-disc in the photograph equalled the same distance as determined perimetrically. (iii) There is a partial defect to the left of the scotoma. This defect, which was detectable only with the dimmest stimulus, is almost certainly caused by damage to underlying fibres from the temporal retina, which traverse the lesion en route to the optic disc.

(*b*) *Changes in fixation.* The procedure for obtaining standard reference photographs was repeated 8 months after retinal operation. Results are shown in Figure 5, centre. The animal maintained the 10° eccentric downward fixation first observed after the cortical operation. As before, the eccentricity was detected during routine perimetry, for the scotoma and natural blind spot lay beneath the fixation mirror.

(4) *Effects of retinal destruction without prior striate cortex damage* (*Animal Scruffy*). (*a*) *Visual field.* A slightly smaller region of the central retina was deliberately destroyed in each eye of this animal. The approximate position of the defect was first determined 2 months later using a $\frac{1}{2}°$ 550 cd/m² stimulus. It was then accurately plotted after reducing the stimulus intensity by 2 log units. A similar stimulus was used to detect the centre of the natural blind spot. Results are shown in Figure 8.

FIGURE 8



Macular scotoma following destruction of central retina, without prior damage to striate cortex. Defect plotted over a 2-month period approximately 3 months after operation. Symbols as in Figure 7. For further details see text.

The symbols correspond to those used for Figure 7. Once again the boundary of the defect is drawn *between* points where stimulus detection is almost totally perfect and where it is almost totally imperfect, or through points where the stimulus is detected about 50 per cent. of the time. At the nine points lying entirely within

the defect there were 11 detected stimuli compared with 81 errors; one detection would be expected by chance after taking into account the overall percentage of false positives when a stimulus of this intensity is presented in intact parts of the visual field. The ratio of detected to undetected stimuli is therefore about 0·1:1, and a defect as severe as this has never been observed after a striate cortex ablation. The ratio at the centre of the natural blind spot was 0·13:1 and it is therefore certain that the macular defect is a true scotoma. The fact that the animal occasionally reports a visual stimulus in a region which should be completely blind is almost certainly accounted for by errors in designating the exact position of the visual axis.

The scotoma appears to be rather smaller than expected from the superimposed outline of the area of retinal destruction (dotted line in Fig. 8), much more so than for the previous animal (see Fig. 7). A likely explanation is that the second animal was much more adept at detecting stray light from a stimulus near the edge of the defect, and indeed it has already been mentioned under preoperative testing that in this animal the real size of the natural blind spot is about 2°–3° larger than indicated by perimetry. The blind spot was therefore mapped accurately with the same stimulus used to explore the macular scotoma and the earlier result was confirmed. The natural blind spot appeared to be about 4° by 3° instead of 7° by 5°. If the appropriate correction is made the macular scotoma fits the outline of the retinal lesion much more closely, perhaps even crossing it at some points.

At the end of testing a 1° 5500 cd./m² stimulus was presented in the centre of the scotoma. There were no detections and 17 errors. This stimulus is usually detected following a striate cortex lesion (Cowey and Weiskrantz, 1963, and earlier sections of this paper), further confirming that cortical and retinal lesions produce qualitatively different field defects.

(*b*) *Changes in fixation.* Together with the eye photographs, the positions of the macular scotoma and natural blind spot with respect to the fixation mirror indicated that throughout post-operative testing the animal's fovea was directed at a point about 5° above the mirror. At first, i.e. 2 months after operation, this eccentricity of fixation was somewhat variable. Although almost always above the mirror, it was sometimes up to 10° to the right of it. This variability decreased until 2 months later the animal consistently fixated vertically above the mirror. At this point in time the procedure for obtaining standard reference photographs was repeated, using eight different mirror positions, two on the central vertical axis of the perimeter and three each on lines 10° to the left and right. Results are shown in Figure 5, right. It will be seen that 75 per cent. of all fixations lie 5° above the mirror after operation. The figure of 14 per cent. for a point 10° above is also well above that for all other parts of the field.

Discussion

Certain conclusions from this study are clear and need little comment. It has been shown that the natural blind spot can be detected perimetrically in monkeys, and that this can be used as a control in studying other defective parts of the visual field. A further control is provided by a retinal lesion, which causes a scotoma. Using these two controls it is certain that the field defects which follow removal of parts of striate cortex are not absolute, even when an entire occipital lobe is removed (Cowey and Weiskrantz, 1963). The improvement in performance seen after striate cortex damage is not entirely spontaneous, for it occurred in one animal whose first post-operative testing took place several years after operation. However, spontaneous recovery cannot be ruled out, for it was found with this animal that parts of an island defect in the temporal field recovered strikingly in function (comparison

of Figs. 2 and 3). The reduction in size of the defect in this region was much more evident than in the macular region and may not have been entirely the result of practice. Such spontaneous recovery has also been observed in man, but usually much sooner after brain damage (Teuber, Battersby and Bender, 1960).

The observations on eccentric fixation confirm and extend earlier findings (Cowey and Weiskrantz, 1963). It is interesting that in three cases of a macular defect (one reported in the earlier study) the animal chose to fixate eccentrically by lowering or elevating its gaze. Lateral displacement has not been observed, except temporarily in an animal suffering from a hemianopic defect (Cowey and Weiskrantz, 1963). The consistency of the eccentric fixation parallels that reported in man after damage to the macular area (Teuber, Battersby and Bender, 1960).

Other observations require some discussion. The macular field defect caused by removal of lateral striate cortex was initially of about the expected size (see Figs. 2 and 3) but became smaller, especially near the fovea (Fig. 4). The prediction about expected size is based upon electrophysiological maps of the retinal projection to striate cortex (Talbot and Marshall, 1941; Daniel and Whitteridge, 1961). These investigators agree so closely about the details of the projection that there is no reason to doubt the accuracy of the maps. The reduction in size cannot be entirely explained by the animal's improving ability to detect stray light near the edge of the defect, for destruction of the macular retina in the same animal produced a field defect which corresponded extremely closely with the area of destruction. This apparently genuine reduction in size, after striate cortex damage, may of course be related to the further problem of why the field defect is not absolute following this operation. There is probably abundant opportunity for stimulation within the macular region to reach intact cortex or sub-cortex (see Cowey and Weiskrantz, 1963, for discussion). For instance, if lateral conduction occurs in the retina one might expect a stimulus close to the edge of a defect to have greater opportunity to influence an intact part of the projection than a similar stimulus at the centre of the defect, which would have to be conducted over a greater region. The shrinkage of the defect with time may reflect the animal's increasing ability to detect such a laterally conducted stimulus near the edge of the defect. Such a mechanism would also explain why the deficit following striate cortex damage is more severe at the centre than at the periphery (see Figs. 2, 3 and 4). This difference between centre and periphery is also genuine for it did not occur following retinal destruction, except at the very edge of the scotoma (see Figs. 6, 7 and 8). It should be stressed that all the results could equally well be explained by assuming that an animal can use information provided by the retinal projection to superior colliculus, if it is assumed that the fovea has a richer and more effective connection than parafoveal regions. But there is insufficient evidence to decide between the alternatives. It will become possible when one has more information about size of receptive fields in monkey retina (Hubel and Wiesel, 1960) and details about the collicular projection. The available evidence is equivocal (Cowey and Weiskrantz, 1963).

It is a pleasure to thank Dr. L. Weiskrantz and Mr. M. J. Morgan for their helpful criticism of the manuscript. The cortical ablation was carried out in conjunction with Dr. Weiskrantz, who kindly checked and assisted in the removal of tissue. I also wish to thank Mr. D. Canwell, who was entirely responsible for the retinal histology, and Mr. R. Hutchison and Miss S. Hopkins who prepared sections of the brain and photographed the Figures. I am particularly indebted to Mr. L. Fison, who made the retinal lesions and arranged that this could be done at Moorfield's Eye Hospital, and to Mr. P. Watson who photographed the retinae of the anaesthetized animals. The work was supported by United States Public Health Services Grant NB-04800-02 and by Medical Research Council Grant G.965/96/B.

References

Cowey, A. (1963). The basis of a method of perimetry with monkeys. *Quart. J. exp. Psychol.*, **15**, 81–90.

Cowey, A., and Weiskrantz, L. (1963). A perimetric study of visual field defects in monkeys. *Quart. J. exp. Psychol.*, **15**, 91–115.

Cowey, A., and Weiskrantz, L. (1967). A comparison of the effects of inferotemporal and striate cortical lesions on the visual behaviour of rhesus monkeys. *Quart. J. exp. Psychol.*, **19**, 246–53.

Daniel, P. M., and Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *J. Physiol.*, **159**, 203–21.

Hubel, D. H., and Wiesel, T. N. (1960). Receptive fields of optic nerve fibres in the spider monkey. *J. Physiol.*, **154**, 572–80.

Talbot, S. A., and Marshall, W. H. (1941). Physiological studies on neural mechanisms of visual localization and discrimination. *Amer. J. Ophthal.*, **24**, 1255–64.

Teuber, H-L., Battersby, W. S., and Bender, M. B. (1960). *Visual Field Defects after Penetrating Missile Wounds of the Brain.* Cambridge, Mass.: Harvard University Press.

Weiskrantz, L., and Cowey, A. (1967). A comparison of the effects of striate cortex and retinal lesions on visual acuity in the monkey. *Science*, **155**, 104–5.

Manuscript received 13th February, 1967.

A COMPARISON OF THE EFFECTS OF INFEROTEMPORAL AND STRIATE CORTEX LESIONS ON THE VISUAL BEHAVIOUR OF RHESUS MONKEYS

BY

ALAN COWEY and LAWRENCE WEISKRANTZ

From the Psychological Laboratory, University of Cambridge

The effects of bilateral removal of inferotemporal cortex or lateral striate cortex are compared. The former operation impairs visual pattern discrimination learning, without disturbing prompt detection and retrieval of food, visual acuity or visual fields. In contrast, animals in which the macular striate projection area has been removed are significantly superior in tests of visual pattern discrimination learning despite showing impairment of visually guided reaching, visual acuity, and visual fields. The results indicate that the inferotemporal defect is not caused by deficient visual sensitivity or acuity.

Introduction

Bilateral removal of inferotemporal cortex of the brain of rhesus monkeys impairs performance on certain kinds of visual tasks without affecting discrimination in other modalities. For example, acquisition of visual pattern or object discriminations is impaired (Mishkin, 1954; Mishkin and Pribram, 1954; Pribram and Barry, 1956; Wilson, 1957), as is retention unless the animals have been overtrained pre-operatively (Chow and Survis, 1958; Orbach and Fantz, 1958). The animals are deficient in acquisition of object learning-set (Wilson and Mishkin, 1959), in retention of learning sets for planometric forms (Chow, 1954), in non-delayed matching to sample (Buffery, 1964), and in visual categorization (Iversen and Weiskrantz, 1967). Generalization gradients to certain visual stimuli are unusually flat (Butter, Mishkin and Rosvold, 1965), and the inferotemporal animals do not utilize as many features of a multi-dimensional stimulus as do normal animals (Butter, Mishkin and Rosvold, 1965; Butter and Gekorski, 1966).

Several explanations of the inferotemporal defect have been proposed (see Diamond and Chow, 1962 for brief review) and it may well be that several are needed. However, only one is considered here. It has been suggested that the inferotemporal animal suffers from some sort of visual sensory loss, characterized by alterations to the visual fields, e.g. an amblyopia, which prevent the animal from seeing clearly. This hypothesis is supported by the following findings. Inferotemporal animals are selectively impaired on size discriminations if the size differences are small (Mishkin and Hall, 1955), and on form discriminations if small forms are used (Pasik, Pasik, Battersby and Bender, 1958, 1960). The difference threshold for discrimination of square vs. triangle is raised at low intensities of illumination but is unchanged at high intensities (Valciukas and Pasik, 1965). Finally, Mishkin and Weiskrantz (1959), found that the critical flicker fusion frequency was depressed in inferotemporal animals as compared with two control groups, one unoperated and the other having frontal cortex resections. Although the authors plausibly ascribe the deficit to a learning defect it is consistent with the idea of sensory impairment, especially since a further group of animals in which the lateral primary visual projection area was removed were similarly impaired.

Despite the fact that inferotemporal animals are often more impaired on the kinds of test mentioned than animals in which the visual striate cortex is damaged and which are therefore presumed to have field defects, it has been argued that the inferotemporal operation produces rather subtle changes in visual field organization, e.g. an amblyopia, which could lead to greater disturbance than a highly circumscribed and severe visual field defect (Pasik, Pasik, Battersby and Bender, 1960). The fact that in man an amblyopia can cause greater visual difficulties than a clear cut scotoma supports this view (Goldstein, 1942).

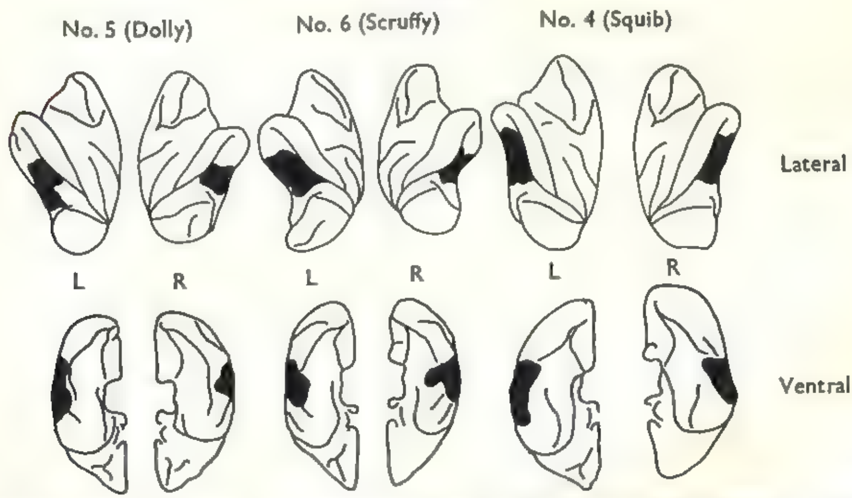
This paper attempts to evaluate the role of possible field defects in causing the inferotemporal syndrome by studying field defects, visual acuity, visually guided reaching, and visual pattern discrimination. The results of perimetric testing for field defects after striate cortex lesions have already been described (Cowey and Weiskrantz, 1963; Cowey, 1967) and are briefly recapitulated only for purposes of comparison. The effects of inferotemporal and striate cortex lesions on acuity have also been described before for all but one of the six animals (Weiskrantz and Cowey, 1963) and can be dealt with swiftly.

Method

Subjects, surgery and histology

Six immature rhesus monkeys, five male, one female, weighing from 4–5 kilos, were used. In three the lateral striate cortex was removed bilaterally, in a further three the inferotemporal cortex was removed bilaterally. Operations were performed by sub-pial suction under deep anaesthesia (Nembutal: sodium pentobarbitone). At the end of the experiment the animals were perfused with normal saline followed by formal-saline and the brains photographed immediately after removal. In the case of the three inferotemporal animals the effects of additional cortical or retinal ablations had been investigated, but these are irrelevant to this paper and for clarity only histological details of the inferotemporal operations are given. Coronal sections, cut at 20 μ, were stained with thionin or cresyl fast violet and used to reconstruct the ablations. Reconstructions for the striate group appear elsewhere (Cowey and Weiskrantz, 1963). Those for the inferotemporal animals are shown in Figure 1.

FIGURE 1



Reconstructions of brain lesions of animals in inferotemporal group. Lateral and ventral views of left and right hemisphere are shown. In every case the lower bank of the middle portion of the inferotemporal sulcus was removed although this is not apparent in outline diagrams. Cross sections are not shown: the inferotemporal ablation is so straightforward to perform that a space-consuming series of cross sectional diagrams for each animal adds little to the surface view.

Reaching tests

The accuracy with which monkeys reached for small pieces of food, on a flat white board measuring 20 in. by 15 in. was observed and recorded on cine film. The board was divided into 1 in. squares by a grid of intersecting black lines, with numbers and letters along two sides so that any square could be identified. Both direction and magnitude of reaching-errors could be recorded by means of the grid. A peanut, raisin, or piece of apple was placed on a square before withdrawing a screen between animal and board. Animals required only a few minutes' experience in this situation before immediately looking and reaching for food when the screen was withdrawn. Food was placed on various squares in random order until, in each testing session, about half the squares had been sampled.

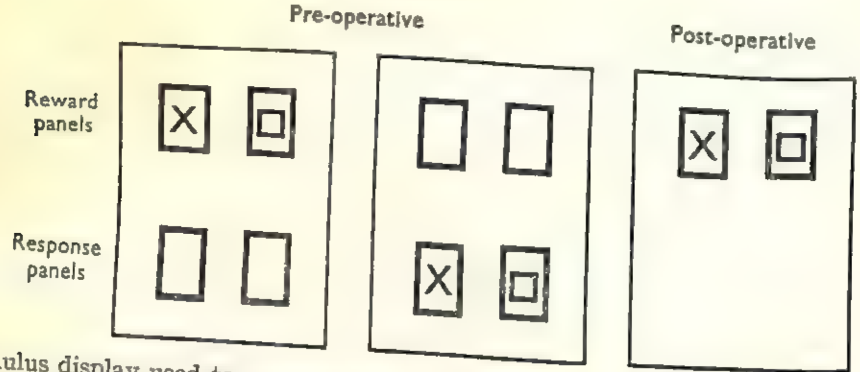
Food was also offered on the experimenter's hand, and occasionally an empty hand was presented to see if an animal would reach for something it could not see.

Visual pattern discrimination

All testing was carried out in a Wisconsin General Test Apparatus, in which the discriminanda were revealed to the animal by lifting an opaque screen. Two tests were given pre-operatively. First, the animals learned to discriminate between two circular cards 3 in. in diameter, each mounted vertically in front of a food cup. One card was plain white, the other consisted of 16 alternating black and white vertical stripes. The animal could obtain food reward by sweeping aside the positive card. The striped card was positive for two of the three animals in each group.

In the second test the animal was presented with a stimulus display consisting of four white panels, two of which carried stimuli. The panels measured 4 in. × 3 in., were 6 in. apart laterally and 10 in. vertically (centre to centre) and were set into a vertical black board (see Fig. 2). The lower pair of panels represented the response site.

FIGURE 2



Stimulus display used to present patterns pre-operatively (left and centre) and post-operatively (right). For explanation see text.

Provided the animal first pushed the positive response panel, the reward panel vertically above it automatically unlocked, permitting the animal to reach for food in a food well behind it. If he first pushed the incorrect response panel, both reward panels remained locked. The stimuli, a black cross vs. a black square, were mounted *either* on the response panels *or* on the reward panels (this being part of a larger experiment on the effects of cue-response and cue-reward spatial separation) and after each animal had mastered the discrimination it was immediately switched to the other condition and retrained to criterion a second time.

Post-operatively the animals were tested for retention of cross vs. square, and for new learning with three pairs of stimuli (see Fig. 2). The vertical stimulus board was used for all post-operative tests but only the two reward panels were used (see Fig. 2, right). The animal could obtain reward by pushing whichever panel bore the positive stimulus. Fifty trials, with correction procedure, were given daily in random order, with a maximum of seven successive identical trials. Criterion was 90 correct responses in 100 successive trials (excluding correction trials).

Visual acuity

Minimum separable visual acuity was measured with diffraction gratings, using a pull-in technique. The apparatus and method, together with results for five of the six animals, have been fully described elsewhere (Weiskrantz and Cowey, 1963).

Visual fields

Field defects were measured in a perimeter designed for use with monkeys. A full description of apparatus and method, together with results for animals with striate cortex lesions has been published (Cowey and Weiskrantz, 1963; Cowey, 1967). The animal was taught to discriminate between a click, and the same click plus a coincident flash of light of variable intensity, size and position. He was rewarded for pressing a lever when the click was presented, and for pressing a different lever when the paired stimulus was used.

Results

Reaching tests

Pre-operatively the animals reached for food correctly and quickly, sometimes before the screen had been fully withdrawn. Misreaching was never observed, although fumbling was seen occasionally, i.e. on less than 1 per cent. of trials. On mock trials, when the experimenter's empty palm or bare testing board were exposed, the animals prepared to reach but never did so.

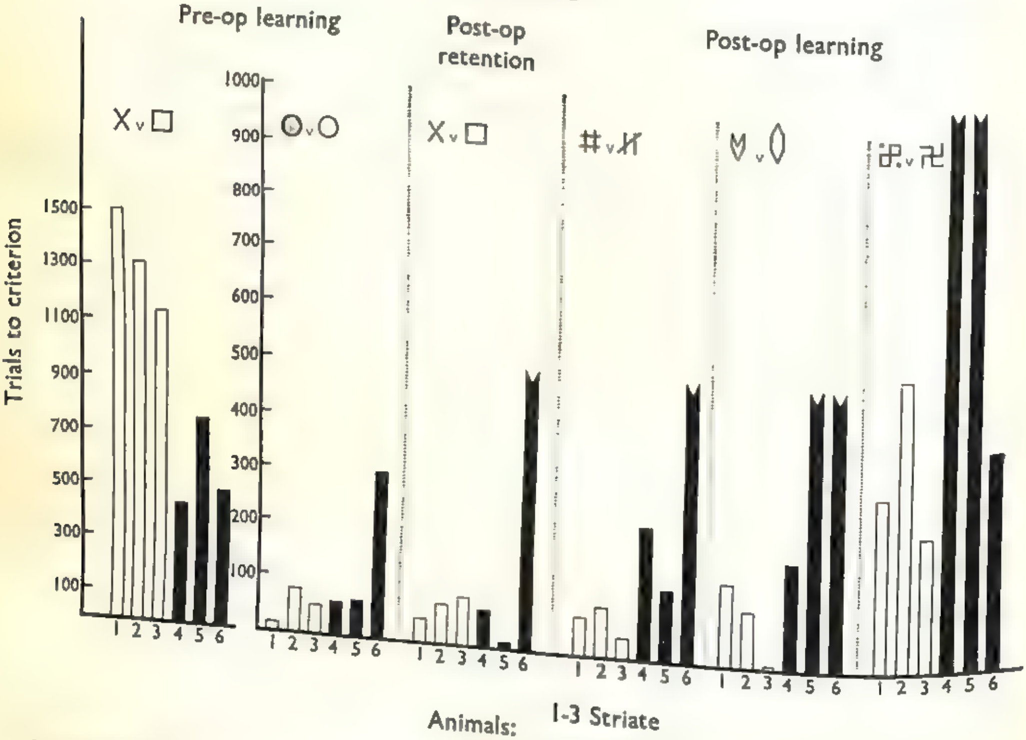
Each animal was tested 24 hours after operation and thereafter daily until performance was perfect on three successive days. All three animals in the striate group showed evidence of imperfect vision, chiefly characterized by misreaching for food. In as many as 20 per cent. of trials in the first 2 days after operation the animals misreached by up to 3 in. Lateral errors were much commoner than over- or under-reaching. Sometimes the animals groped for non-existent food, and attempted to pick up marks on the board or empty peanut skins, which were always ignored pre-operatively. In further contrast to their pre-operative behaviour they frequently reached for and explored the experimenter's empty palm, and chose the incorrect hand when both hands were offered with food on only one of them. The animals peered intently when they failed to locate food, and seemed to be adjusting their gaze. All these symptoms subsided rapidly and disappeared within 10 days. The order in which they did so was not determined quantitatively, but it is our strong impression that accurate reaching for food which contrasted sharply with the background (e.g. raisin) preceded that for a reward such as apple, which was of similar colour to the testing board. The last symptom to disappear was misreaching for a light-coloured reward on the experimenter's hand, probably for the same reason and also because of the confusing contours and interstices compared with the flat board. It should perhaps be stressed that the misreaching referred to here is not attributable to ataxia for (*a*) the animals' movements were otherwise normal, e.g. in conveying food to the mouth; (*b*) reaching was impaired only in direction; there was no tremor and movements were swift; (*c*) it has been shown elsewhere (Cowey and Weiskrantz, 1961) that the recovery from misreaching is postponed if an animal is kept in darkness with no opportunity for visually guided reaching but ample opportunity to recover from a transient ataxia, and (*d*) similar misreaching is seen after destruction of the macular region of the retina, which produces a field defect without damaging any part of the motor system.

The behaviour of the three inferotemporal animals was strikingly different. At no time did they show evidence of defective vision. Pre-operative and post-operative performances were indistinguishable.

Visual pattern discrimination

Results are shown in Figure 3. The pre-operative scores for the cross vs square discrimination, which was presented to each animal in two ways, have been added to give an overall score for each animal. The three inferotemporal animals tended to be superior pre-operatively ($p = 0{\cdot}1$, Mann Whitney test, two-tailed) and this was not an artefact of the order in which the two discrimination conditions.were presented, for the orders were balanced between the groups. This was a fortuitous result, for the animals were selected for operation according to performance on acuity tests and not as a result of scores on pattern discrimination. The groups were clearly not different on stripes vs. plain discrimination.

FIGURE 3



Histograms of visual discrimination performance before and after operation. Performance of animals in inferotemporal group is shown in solid black. For each problem the number of trials taken to reach the beginning of criterion (90 correct out of 100 successive trials) is given. Indentation at the top of any bar indicates that the animal failed to reach criterion and was stopped at that point. Note different ordinate scales for the first and for all subsequent problems.

Post-operatively the two groups were not significantly different on the retention test ($p = 0{\cdot}65$, Mann Whitney test, one-tailed), which is hardly surprising since the animals had twice reached criterion pre-operatively and overtraining is known to abolish or reduce the retention deficits caused by inferotemporal ablation (Chow and Survis, 1959; Orbach and Fantz, 1958). Despite this, one inferotemporal animal failed to reach criterion in 500 trials.

The two groups differed significantly in tests of post-operative learning. For each of the three problems the statistic U (Mann Whitney test, one-tailed) was equal

to or less than 1. Since in each case the probability that $U \leqslant 1$ is 0·1, the probability that this will be so in every case is 0·001. It is concluded *that the inferotemporal animals are significantly impaired on acquisition of visual pattern discriminations* as compared with striate controls.

Visual acuity

The effects of the operations on the visual acuity of five of the six animals have been described elsewhere (Weiskrantz and Cowey, 1963). They are reproduced in Table I, together with data for a third inferotemporal subject (number 6). Only

TABLE I

MINIMUM SEPARABLE VISUAL ACUITY IN MINUTES OF ARC

	Striate			*Inferotemporal*		
	1	*2*	*3*	*4*	*5*	*6*
Pre-op.	0·57	0·98	0·72	0·60	0·58	0·70
Post-op.	0·64	0·94	0·92	0·58	0·52	0·69
$\frac{\text{Pre}}{\text{Post}} \times$ 100 per cent.	89	104	78	103	111	101

one striate animal, in which a small part of the foveal projection area was spared in one hemisphere, was not impaired. No inferotemporal was impaired; in fact all three showed a slight improvement. However, animal number 6 initially performed at chance after operation and required lengthy retraining at stimulus angles well above final threshold. He behaved as if he had not seen the acuity stimuli before, although comprehension of the testing situation was perfect.

Visual fields

All three animals in the striate group had visual field defects as tested perimetrically (Cowey and Weiskrantz, 1963). By contrast, there was no evidence of a field defect in any inferotemporal animal, even with a $\frac{1}{2}°$ 5·5 cd./m^2 stimulus which was used to plot the natural blind spot in animal number 6 (Cowey, 1967). However, one inferotemporal animal (number 4) performed at chance levels in the perimeter for the first weeks after operation before suddenly and perfectly re-acquiring the discrimination between a click and click-plus-light. Since the animal reported a non-existent visual stimulus as often as it failed to report a real flash of light the temporary defect appeared not to be caused by inability to see the visual stimulus. This was confirmed by presenting light flashes without the usual coincident auditory stimulus. The animal always responded although not necessarily by pressing the correct lever.

DISCUSSION

It has been shown that bilateral removal of inferotemporal cortex produces significantly greater impairment of visual discrimination learning than ablation of the macular projection area, without disturbing visually guided reaching, visual acuity, or visual fields, all of which are adversely affected by damage to the primary

visual projection area. This result clearly fails to support the notion that the inferotemporal operation produces "an alteration in the visual field (an amblyopia)"—Pasik, Pasik, Battersby and Bender (1960). If the amblyopia is so subtle that it leaves acuity intact and altogether escapes detection in a perimeter it is difficult to see why it should affect the perception of prominent patterns. At present, other explanations of the inferotemporal defect, in terms of memory loss (Iversen and Weiskrantz, 1964) or selective attention to particular stimulus dimensions (Butter and Gekorski, 1966), seem more convincing.

We wish to thank Mr. R. Hutchison and Miss S. Hopkins for doing the histology and photographing the Figures, and Mr. D. Nicholas for his help in testing the animals.

The work was supported by United States Public Health Service Grant NB-04800-02 and by Medical Research Council Grant G.965/96/b.

References

Buffery, A. W. H. (1964). *The Effects of Frontal and Temporal Lobe Lesions upon the Behaviour of Baboons.* Ph.D. Thesis, University of Cambridge.

Butter, C. M., and Gekorski, W. L. (1966). Alterations in pattern equivalence following inferotemporal and lateral striate lesions in rhesus monkeys. *J. comp. physiol. Psychol.*, **61**, 309–12.

Butter, C. M., Mishkin, M., and Rosvold, H. E. (1965). Stimulus generalization in monkeys with inferotemporal and lateral occipital lesions. In Mostofsky, D. J. (Ed.), *Stimulus Generalization.* Stanford: Stanford University Press.

Chow, K. L. (1954). Effects of temporal and neocortical ablation on visual discrimination learning sets in monkeys. *J. comp. physiol. Psychol.*, **47**, 194–8.

Chow, K. L., and Survis, J. (1958). Retention of overlearned visual habit after cortical ablation in monkey. *Arch. Neurol. Psychiat.*, **79**, 640–6.

Cowey, A. (1967). Perimetric study of visual field defects in monkeys after striate cortex and retinal ablations. *Quart. J. exp. Psychol.*, **19**, 232–45.

Cowey, A., and Weiskrantz, L. (1961). Role of experience in misreaching caused by striate cortex lesions. *Nature*, **192**, 1319.

Cowey, A., and Weiskrantz, L. (1963). A perimetric study of visual field defects in monkeys. *Quart. J. exp. Psychol.*, **15**, 91–115.

Diamond, I. T., and Chow, K. L. (1962). Biological psychology. In Koch, S. (Ed.), *Psychology: A Study of a Science*, Vol. IV. New York: McGraw-Hill.

Goldstein, K. (1942). *After Effects of Brain Injuries in War.* New York: Greene and Stratton.

Iversen, S. D., and Weiskrantz, L. (1964). Temporal lobe lesions and memory in the monkey. *Nature*, **201**, 740–2.

Iversen, S. D., and Weiskrantz, L. (1967). Perception of redundant cues by monkeys with inferotemporal lesions. *Nature*, **214**, 241–3.

Mishkin, M. (1954). Visual discrimination performance following partial ablations of the temporal lobe. II. Ventral surface vs. hippocampus. *J. comp. physiol. Psychol.*, **47**, 187–93.

Mishkin, M., and Hall, M. (1955). Discrimination along a size continuum following ablation of the inferior temporal convexity in monkeys. *J. comp. physiol. Psychol.*, **48**, 97–101.

Mishkin, M., and Pribram, K. H. (1954). Visual discrimination performance following partial ablations of the temporal lobe. I. Ventral vs. lateral. *J. comp. physiol. Psychol.*, **47**, 14–20.

Mishkin, M., and Weiskrantz, L. (1959). Effects of cortical lesions in monkeys on flicker fusion frequency. *J. comp. physiol. Psychol.*, **52**, 660–6.

Orbach, J., and Fantz, R. L. (1958). Differential effects of temporal neocortical resections on overtrained and non-overtrained visual habits in monkeys. *J. comp. physiol. Psychol.*, **51**, 126–9.

Pasik, P., Pasik, T., Battersby, W. S., and Bender, M. B. (1958). Visual and tactual discriminations by macaques with serial temporal and parietal lesions. *J. comp. physiol. Psychol.*, **51**, 427–36.

Pasik, T., Pasik, P., Battersby, W. S., and Bender, M. B. (1960). Factors influencing visual behavior of monkeys with temporal lobe lesions. *J. comp. Neurol.*, **115**, 89–102.

PRIBRAM, H. B., and BARRY, J. (1956). Further behavioral analysis of parieto-temporo-preoccipital cortex. *J. Neurophysiol.*, **19**, 99–106.

VALCIUKAS, J. A., and PASIK, P. (1965). Influence of illumination on pattern discrimination thresholds in monkeys with temporal lobe lesions. *Fed. Proc.*, **24**, No. 2, Part 1, 274.

WEISKRANTZ, L., and COWEY, A. (1963). Striate cortex lesions and visual acuity of the rhesus monkey. *J. comp. physiol. Psychol.*, **56**, 225–31.

WILSON, M. (1957). Effects of circumscribed cortical lesions upon somesthetic and visual discrimination in the monkey. *J. comp. physiol. Psychol.*, **50**, 630–5.

WILSON, W. A., and MISHKIN, M. (1959). Comparison of the effects of inferotemporal and lateral occipital lesions on visually guided behavior in monkeys. *J. comp. physiol. Psychol.*, **52**, 10–7.

Manuscript received 13th February, 1967.

EXPLORATORY BEHAVIOUR IN ELEVATED AND ENCLOSED MAZES

BY

M. S. HALLIDAY

From the Laboratory of Experimental Psychology, University of Sussex

Rats given daily exploratory trials in a simple enclosed maze showed a decline in activity from day to day, while animals exploring an elevated maze showed no decline between days. In a further experiment exploration of either an elevated or an enclosed maze was found to cease after about 15 min. The implications of these findings are discussed.

Introduction

Decrements in rats' exploratory activity have commonly been found within a single exploratory session (e.g. Montgomery, 1951; Berlyne, 1955). Various theoretical explanations have been given for this intra-session decrement (Montgomery, 1953*a*; Glanzer, 1953; Berlyne, 1950). There has, however, been disagreement as to whether *inter*-session decrements in exploratory behaviour occur in simple mazes. Montgomery claimed that no such decrements occurred between daily trials, but his own results do not entirely bear out this contention. In three of his experiments (Montgomery, 1953*b*, 1953*c*; Montgomery and Monkman, 1955), in which trials were given on two successive days in simple enclosed mazes, a decline in activity did occur between the first and second days; it is impossible to say from the results reported by Montgomery whether these declines were statistically significant, but in two cases they were substantial. Evidence of inter-session declines in exploratory activity in enclosed Y-mazes is also reported by Carr and Williams (1957), Williams and Kuchta (1957) and Ehrlich (1959). But in elevated mazes Montgomery found slight inter-session increases (Montgomery, 1952); inter-session increases in elevated mazes have also been reported by Steinberg, Rushton and Tinson (1961) and Watson (1964).

The present experiments attempt to reconcile these results by comparing directly exploratory behaviour in elevated and enclosed mazes. Such a comparison is also of interest since it is generally accepted that elevated mazes are more "fear-evoking" than enclosed ones (Montgomery, 1955). Thus the comparison of exploratory activity in the two types of maze should throw light on the relationship between exploratory behaviour and fear.

Experiment IA

Subjects

The subjects were 24 naive, female hooded rats obtained from the Medical Research Council, Mill Hill, London. They were about 100 days old at the time of the experiment. The rats were housed, four to a cage, in wire mesh cages and had *ad lib* access to food (laboratory chow) and water.

Apparatus

Two mazes were used: (*a*) *Elevated Y-maze.* An elevated maze with three arms meeting at 120°; each of these arms was 18 in. long, 4 in. wide and 18 in. above the floor level. (*b*) *Enclosed Y-maze.* This was exactly similar in plan and dimensions to the elevated maze, but had 6 in. high unpainted metal walls and removable transparent Perspex roofs to the alleys. When in use these mazes were placed in an enclosure, 4 ft. 6 in. × 4 ft. 6 in. × 4 ft. high, which was covered with black curtain material. The

enclosure was lighted by a single 100 watt bulb, suspended centrally 6 in. from the top. The enclosed maze was placed on blocks so as to bring it up to the same height (and distance from the light) as the elevated maze. The rats could be observed through a screened aperture in one side of the enclosure. A white noise of moderate intensity masked sounds from outside the experimental room.

Procedure

All subjects were given 3 days pre-training in which they were allowed to move around on a table for about 5 min., being picked up and handled from time to time.

Twelve subjects were randomly assigned to each of two groups (*El* and *En*). Rats in both groups received one daily 3-min. trial for 6 days. On each trial the rat was put in the maze at the end of the same arm, allowed to move around for 3 min., and then removed. Rats in Group El explored the elevated maze on each day; rats in Group En explored the enclosed maze. All testing was carried out between 2.00 p.m. and 4.00 p.m.

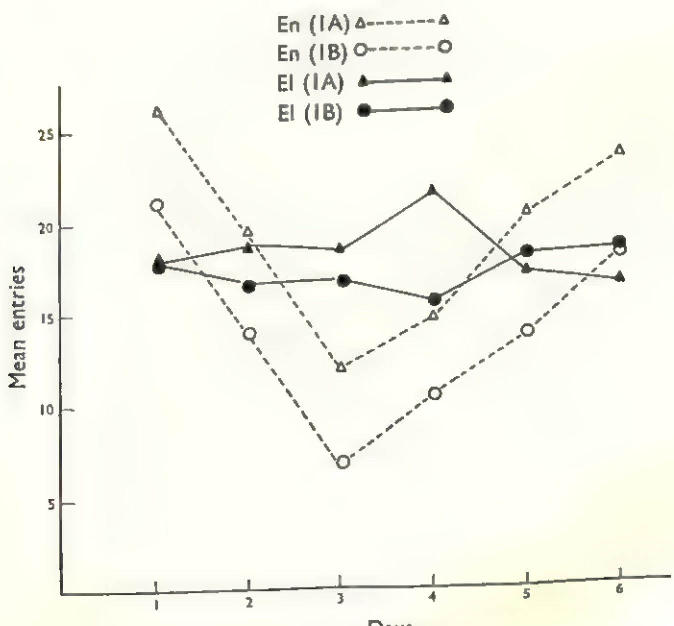
Each arm of the maze was divided into two 9 in. sections by a line on the floor, and the entries to each of these sections during each ½ min. of the trial were recorded on a cyclostyled sheet. A rat was considered to have entered a maze section if its head and three of its feet were in the section. The number of fecal boli deposited by each rat on each trial was also recorded.

Results

The mean number of section entries per day for each group is shown in Figure 1. Table I gives the mean daily scores and also the percentage alternation on each trial; that is the percentage of occasions on which, when arriving at the choice point of the maze, the rat entered the arm it visited least recently.

The activity of the subjects in Group El varied considerably and unsystematically from day to day; however, a Friedman two-way analysis of variance showed no significant variation between days, nor were the differences between any two successive daily scores significant. The changes in daily activity scores for the

FIGURE 1



Mean daily activity scores in the elevated and enclosed mazes. Experiments IA and IB.

TABLE I

MEAN EXPLORATORY ACTIVITY AND PERCENTAGE ALTERNATION IN ELEVATED AND ENCLOSED MAZES. (EXPERIMENT IA)

	Day					
	1	2	3	4	5	6
Mean activity:						
Group El ..	17·9	18·7	18·35	21·3	16·8	16·3
Group En ..	26·2	19·3	11·65	14·4	20·0	23·2
Percentage alternation:						
Group El ..	66·2	61·5	59·5	62·5	69·2	76·5
Group En ..	68·3	68·0	78·0	79·0	69·5	66·7

subjects in Group En were significant ($p < 0·05$, Friedman). This was due to the fall and subsequent rise in daily scores. The decrease in activity scores between Day 1 and Day 3 and the subsequent rise between Day 3 and Days 5 or 6 were each significant ($p < 0·012$, Signs test). The rats in Group En were significantly more active on Day 1 than those in Group El ($p < 0·02$, U-test) but they were significantly less active than Group El on Days 3 and 4 ($p < 0·02$, U-test). (All tests of significance in this paper are two-tailed.)

There was no systematic change in alternation scores with successive trials in either group; though Group El alternated (non-significantly) less than Group En.

In Group El the number of rats which defecated remained fairly constant (Mean: Days 1–3 = 7·3; Days 4–6 = 6·7). In Group En the number declined from five on Day 1 to one on Day 6; the negative correlation between trial number and number of animals defecating was significant ($p < 0·012$, Kendall's τ). The rats in Group El defecated significantly more over the period as a whole than those in Group En ($p = 0·002$, Signs test); and on Days 3 and 4, the days when Group El was significantly more active than Group En, the rats in Group El defecated significantly more than those in Group En ($p < 0·05$).

EXPERIMENT IB

In view of the rather surprising results of the previous experiment and the variability in activity scores in the elevated maze the experiment was replicated, after a year's interval, with further groups of 12 rats in each maze. Subjects, apparatus and procedure were similar to those in Experiment IA except that the maze arms were 1½ in. longer and the trials were given between 10.00 a.m. and 12.00 noon.

Results

The daily activity and alternation scores are summarized in Table II and Figure 1.

The activity of the animals in Group El remained constant from day to day (no significant variation by the Friedman two-way analysis of variance).

The changes in daily activity scores of the animals in Group En were significant ($p < 0·05$, Friedman). This was due to the fall and subsequent rise in daily scores. The decline in activity scores between Day 1 and Day 3 and the subsequent rise between Day 3 and Days 5 or 6 were significant in each case ($p < 0·006$, Signs test). The only significant difference in activity between groups occurred on Day 3 when Group El was more active than Group En ($p < 0·01$, U-test).

TABLE II

Mean Exploratory Activity and Percentage Alternation in Elevated and Enclosed Mazes. (Experiment IB)

	Day					
	1	2	3	4	5	6
Mean activity:						
Group El ..	17·8	16·5	16·6	15·3	17·8	18·1
Group En ..	21·0	13·75	6·7	10·2	13·4	17·9
Percentage alternation:						
Group El ..	76·8	62·0	72·5	60·0	61·4	63·6
Group En ..	76·1	74·4	89·5	64·5	71·8	70·0

The number of rats defecating in Group El remained more or less constant with successive daily trials; 10 rats defecated on Day 1 and nine on Day 6; on only one day (Day 4) did as few as eight animals defecate. In Group En there was once again a negative correlation between number of animals defecating and trial number ($p = 0{\cdot}05$, Kendall's τ). Over the period of the experiment as a whole the rats in Group El defecated significantly more than those in Group En ($p = 0{\cdot}02$, U-test).

The results of the replication confirm those of Experiment IA. In this experiment the scores in the elevated maze were less variable than in Experiment IA, but in terms of the statistical analysis they were similar. The subjects in the enclosed maze were consistently more active in this experiment than in Experiment IA, but the daily changes in activity were highly similar. The results of the two experiments may therefore conveniently be discussed together.

Discussion

It is apparent that there was no appreciable inter-session decline in exploratory activity in the elevated Y-maze, while there was such a decline, followed by a recovery, in the enclosed maze. This result resolves the discrepancies in the experimental literature already discussed, since in these cases too there were inter-session declines in enclosed but not in elevated mazes. The increase in activity in the enclosed Group after Day 3 is a curious effect which has not previously been reported; Williams and Kuchta's (1957) experiment is the only directly comparable one in which trials were continued for more than 4 days. Such an increase is unexpected and is not in accord with any theory of exploration.

The results for the first 3 days present difficulties for any theory of exploration which supposes that novelty of the environment is the sole determinant of exploratory behaviour; for it appears that the enclosed maze becomes less novel with repeated exposure while the elevated maze does not. It might be that the elevated maze offered more to explore than the enclosed one and so the rats did not appreciably reduce its novelty during any one session. If this were so, rats would be expected to continue to explore an elevated maze for longer than an enclosed maze during a single prolonged exposure. A further experiment was therefore undertaken to investigate this possibility.

Experiment II

Subjects, apparatus and procedure

The subjects were 32 naive, female, hooded rats, obtained from the Medical Research Council, Mill Hill, London. They were about 100 days old. The same apparatus was used as in Experiment IB.

The subjects were housed, fed and pretrained exactly as in the two previous experiments.

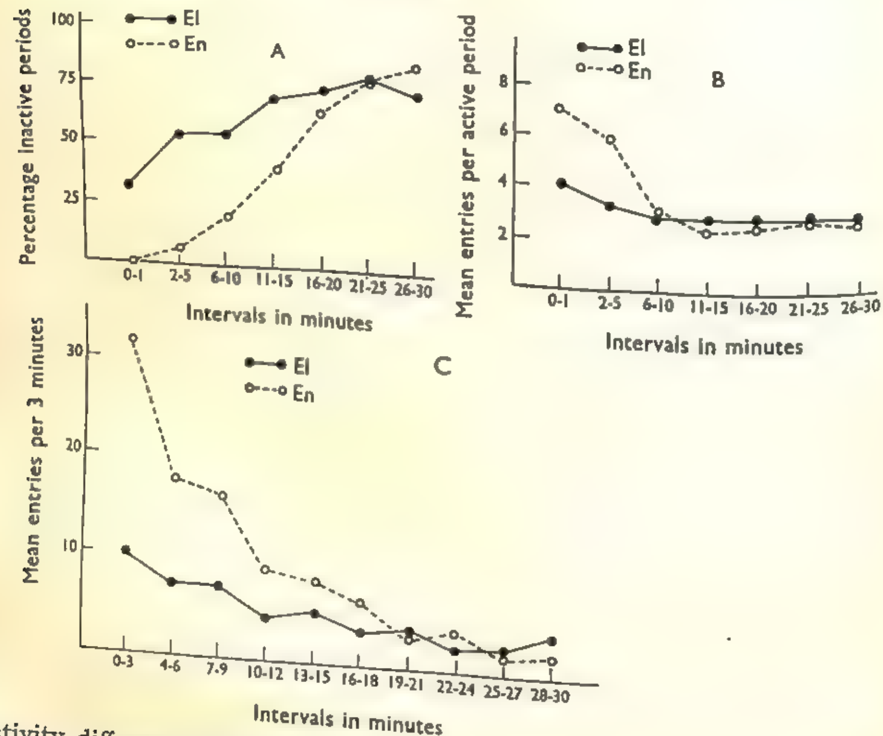
On the day following the end of pretraining all subjects were given a single 30-min. trial. Twenty-two rats were tested in the elevated maze and 10 in the enclosed maze. Testing and recording procedures were as in the two previous experiments.

Results

Elevated

Figure 2C shows the mean activity of the rats in the elevated maze for each 3 min. of the trial; there was a steady fall-off in activity up to about 15 min. There were significant decreases in activity between both the first and the second, and the second and the third 5-min. periods of the trial (both $p < 0{\cdot}05$, Wilcoxon). There was also a significant difference between the second and either the fourth or the fifth such period ($p < 0{\cdot}04$, Signs test). There were no significant differences in amounts of activity between the last three 5-min. periods.

FIGURE 2



Activity differences during a single 30-min. trial in elevated and enclosed mazes. Experiment II. A. Percentage of Inactive Periods. B. Mean Entries per Active Period. C. Mean Entries.

A more detailed analysis of the activity scores was also made. For recording purposes a trial was divided into 60 30-sec. intervals. An "active period" means any one of these 30-sec. intervals during which the animals made two or more entries; an "inactive period" means any such interval in which no entries were made. Table III and Figures 2A and 2B show for the first min., the second to fifth min. and each succeeding 5 min. of the trial, the mean number of entries in each "active period" and the percentage of 30-sec. periods on which the rats were inactive.

TABLE III

MEAN ENTRIES IN ACTIVE PERIODS, AND PERCENTAGE OF INACTIVE PERIODS

	Time intervals in min.						
	0–1	*2–5*	*6–10*	*11–15*	*16–20*	*21–25*	*26–30*
Group El:							
Mean activity per active period ..	4·18	3·50	3·16	3·23	3·225	3·24	3·03
Percentage inactive periods	36·4	54·0	55·9	69·0	74·5	77·7	73·2
Group En:							
Mean activity per active period ..	7·25†	6·1*	3·35	2·65	2·71	2·94	2·92
Percentage inactive periods	0*	6·25†	20·0*	37·3*	62·5	78·7	83·2

The mean entries per active period is the sum, across all subjects, of the activity scores during active periods, divided by the number of active periods.

* Difference between Group El and Group En significant at 0·01 level or better.

† Difference between Group El and Group En significant at 0·005 level or better. U-tests were used for mean activity scores, and χ^2 tests for inactive period scores.

The mean activity in active periods was constant after the first 5 min. Activity per active period decreased significantly between the first and second 5-min. period ($p = 0{\cdot}04$, Signs test) but this was mainly due to very high scores in the first min.; the difference between the mean activity per active period in the 2–5-min. interval and the 6–10-min. interval was non-significant.

The "percent inactive" scores rose steadily up to the fifth 5-min. period. The only difference between scores on successive 5-min. periods that approached significance was between the second and third 5-min. periods ($p = 0{\cdot}076$, Signs test); however the scores in the first 10 min. and the last 10 min. differed at the 0·001 level (Signs test).

Table IV shows the percentage alternation for each 5-min. period of the trial, together with the significance, by the Signs test, of the deviations from 50 per cent. (chance alternation). The scores from 11–30 min. combined were at the level of chance (94 alternations out of 187 choices). The decline in alternation between 0–10 min. and 11–30 min. was significant ($p = 0{\cdot}004$, Signs test).

TABLE IV

PERCENTAGE ALTERNATION AND SIGNIFICANCE OF DEVIATION FROM CHANCE

	Time intervals in min.					
	0–5	*6–10*	*11–15*	*26–20*	*21–25*	*26–30*
Group El:						
Percentage alternation ..	67·7	62·5	49·25	52·2	59·5	42·9
Significance level (p) ..	0·003	0·046	n.s.	n.s.	n.s.	n.s.
Group En:						
Percentage alternation ..	64·1	61·2	42·0	47·4	50·0	50·0
Significance level (p) ..	0·001	0·045	0·215	n.s.	n.s.	n.s.

Enclosed

The changes in activity during the period of exploration are shown in Figure 2. Once again there was a fall in activity up to about 15 min. followed by a low but constant amount of exploration. The decrease in activity between the first and second min. was significant ($p = 0{\cdot}04$, Signs test), as were the decreases between the first and second ($p = 0{\cdot}002$, Signs test), and the second and third 5-min. periods ($p = 0{\cdot}01$, Wilcoxon). The differences between the third and fifth, and third and sixth, 5-min. periods were also significant at below the 0·05 level (Signs test), but no differences between the last three 5-min. periods were significant.

An analysis of the results in terms of active periods and percentage inactive is given in Table III. The activity per active period scores were fairly constant from 5 min. onwards and stabilized at much the same level as those for the elevated maze. The decrease in activity per active period was significant between the first min. and second to fifth min., and also between the first and second 5-min. periods ($p < 0{\cdot}001$ in each case, Signs test); no further changes were significant. The increase in percentage of inactive periods was steady, though no two adjacent 5 min. scores differed significantly; but the difference between the first and last 10-min. periods was significant at below the 0·001 level (Signs test).

The alternation scores for the enclosed maze are shown in Table IV together with the significance of their deviation from chance alternation. The decline in alternation scores between 0–10 min. and 11–30 min. was significant ($p = 0{\cdot}02$, Signs test). As in the elevated maze the alternation scores from 11–30 min. were at the chance level.

Comparison of groups

It can be seen from Figure 2 and Table III that the enclosed animals were more active for at least the first 10 min. The significances of the differences in these scores are given in the Tables. The differences in activity scores in the first 5 min. arose both from greater activity within 30 sec. active periods and from a lower proportion of inactive periods; between 5 and 15 min. the difference was entirely due to differences in proportion of inactive periods.

The alternation scores for both groups were very similar both in amount of alternation and in change with time. In both cases alternation dropped to chance at about the time at which activity reached a steady low level.

Discussion

In both mazes exploratory activity decreased with time in the maze, and in both cases it reached an asymptotic low level after 11–15 min. After this the rats' activity was probably largely non-exploratory in nature (not primarily controlled by the stimuli in the environment), since alternation scores dropped to chance.

The analysis of the results also shows that decrements in exploratory activity occurred, not because the rats were moving more slowly, but because they were moving less often. If a rat was active at all it was likely to be as active at the end as at the beginning of the trial; it was the probability of activity which changed with time. Typically a rat would run from the end of one arm of the maze to the end of another arm (thus scoring three section entries) and then pause before making another "run." If it was very active, and the pause between "runs" short, two or more such "runs" could occur within one 30 sec. recording period thus raising the activity per active period scores; this appeared to happen in the first min. in the elevated maze and the first 5 min. in the enclosed maze. This did not, however, represent a

discontinuity in behaviour, which can best be thought of as bursts of activity of constant amplitude but decreasing frequency. In this respect the pattern of behaviour is similar to that found in the extinction of a learned response (Hurwitz, 1957).

The rats' activity in the two types of maze, in this single exposure, differed because the rats in the elevated maze made longer pauses between active periods than those in the enclosed mazes. But when they were moving around both groups were equally active. Of the various possible explanations for this difference the most likely seems to be that exploration of the two types of maze proceeded rather differently. In the enclosed maze by far the greater part of the stimuli to explore were within the maze itself (the view through the roof was very limited); the only way to explore such stimuli was to approach and inspect them. In the elevated maze the extra-maze stimuli were relatively far more important, and these stimuli could, to some degree, be investigated without moving around. Qualitative observation of the subjects' behaviour supported this view. In the enclosed maze, when not moving around the rats preened or were almost immobile. In the elevated maze they reared up on their hind legs, moved their heads around, or peered over the edges of the maze; they were not scoring any entries, but they were clearly exploring; the frequency of this type of behaviour tended to decrease as the trial continued.

The results indicate that rats stop exploring elevated and enclosed mazes after about the same period of time, and this suggests that elevated mazes are not more interesting than enclosed mazes. Taken in combination with the results of the previous experiments in which it was found that there was an inter-session decline with daily trials in the enclosed maze but not in the elevated maze these results pose a problem in interpreting the difference between exploration in the two types of maze. This difficulty is particularly acute for any theory of exploration which tries to account for the results of both experiments solely in terms of what the rat learns about the mazes. Such a theory must explain the results of both experiments in terms of the decreasing novelty of the mazes; but the only difference there can be between the two situations is that, in the case of the daily trials the novelty of the maze may be affected by how much the rats forget in the intervals between trials, while in the single long exposure such forgetting is relatively unimportant. Such differences might well affect the rate at which the maze became familiar and so produce different rates of change of exploratory activity in the two experiments. But these differences in forgetting could not alone account for the different relationship between elevated and enclosed groups in the two experiments.

There must therefore be some further difference between the two types of maze which influences their capacity to elicit exploratory behaviour. That is, experience of the maze must have a different effect on the animal in the two types of situation, such that familiarity with the maze changes the rats' response to the situation during daily trials in the enclosed, but not in the elevated maze. The results on defecation in Experiments IA and IB suggest one way in which the animals' reaction to the mazes changed from day to day in just this manner. It is generally accepted that, as a group measure, defecation is an indicator of fear or emotionality in rats (e.g. Broadhurst, 1957). It thus appears from the defecation scores that, in Experiments IA and IB, the rats in the enclosed maze became less fearful with repeated experiences of the maze, while the elevated maze evoked fear to about the same degree on all days. Observation of the animals' general behaviour in the two mazes accorded with this conclusion.

It appears, therefore, that the rats' fear of the two types of maze changed in very much the same way as their exploratory behaviour, over the first 3 or 4 days.

Similarly it is reasonable to suppose that fear of both types of maze habituated steadily during the single long exposure described in Experiment II. (The proportion of animals defecating in any one 5-min. period decreased steadily throughout the trial in both mazes, but this might only represent a decreased capacity to defecate). In this case too the rats' fear of the mazes probably varied in much the same way as their exploratory activity. Given that some factor other than novelty is needed to explain the results on exploratory activity, the findings on defecation make it seem possible that fear is of greater importance in the control of this sort of behaviour than has commonly been supposed.

It is tentatively suggested, therefore, that in some cases exploratory behaviour in rats is evoked by the capacity of stimuli to evoke fear, rather than by the novelty of the stimuli. Strongly aversive stimuli will, of course, be avoided, but stimuli which evoke milder degrees of fear may be approached and investigated; this investigation will either reduce the animal's fear of the stimulus to the point where it stops investigating, or alternatively, if the stimulus turns out on investigation to be highly aversive, and animal will come to avoid rather than explore it. Most novel stimuli arouse fear, and if the above suggestions are correct, will therefore elicit exploratory behaviour; in general such stimuli will, as they become familiar, also lose their capacity to evoke fear, but some types of environment remain fear-evoking even after they have become familar. It is suggested, on the basis of the defecation results, that the elevated maze is an environment of the latter sort. Within any one experimental session fear of the maze will habituate and exploratory activity decrease, but by the next trial the maze will have recovered its capacity to arouse fear and hence exploratory activity, although it may be more familiar on this occasion than on the first trial. The enclosed maze differs from the elevated one in that it loses its capacity to evoke fear as it becomes more familiar; hence there are both inter- and intra-session declines in activity in the enclosed maze.

This explanation of the results is admittedly speculative but it has firm factual foundations. (1) The results are not explicable in terms of novelty and familiarity of the mazes, some other variable is required. (2) The fear-evoking properties of the mazes do appear to change in just the way required by the theory. (3) Further experiments, in which the length of trials and delays between trials have been systematically varied have confirmed the systematic nature of the differences in exploratory activity in elevated and enclosed mazes, and have given further support to the theoretical framework outlined above.

This work was carried out as part of a programme supported by the Medical Council. I am grateful to Mr. A. J. Watson for his help and advice.

References

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.*, **41**, 68–80.

Berlyne, D. E. (1955). The arousal and satiation of perceptual curiosity in the rat. *J. comp. physiol. Psychol.*, **48**, 238–46.

Broadhurst, P. L. (1957). Determinants of emotionality in the rat. I. Situational factors. *Brit. J. Psychol.*, **48**, 1–12.

Carr, R. M., and Williams, C. D. (1957). Exploratory behaviour of three strains of rats. *J. comp. physiol. Psychol.*, **50**, 621–3.

Ehrlich, A. (1959). Effects of past experience on exploratory behaviour in rats. *Canad. J. Psychol.*, **13**, 248–54.

Glanzer, M. (1953). Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.*, **60**, 257–68.

Hurwitz, H. M. B. (1957). Periodicity of response in operant extinction. *Quart. J. exp. Psychol.*, **9**, 177–84.

MONTGOMERY, K. C. (1951). The relation between exploratory behaviour and spontaneous alternation in the white rat. *J. comp. physiol. Psychol.*, **44**, 582–9.

MONTGOMERY, K. C. (1952). Exploratory behaviour and its relation to spontaneous alternation in a series of maze exposures. *J. comp. physiol. Psychol.*, **45**, 50–7.

MONTGOMERY, K. C. (1953*a*). Exploratory behaviour as a function of "Similarity" of stimulus situations. *J. comp. physiol. Psychol.*, **46**, 129–33.

MONTGOMERY, K. C. (1953*b*). The effect of activity deprivation on exploratory behaviour. *J. comp. physiol. Psychol.*, **46**, 438–41.

MONTGOMERY, K. C. (1953*c*). The effect of the hunger and thirst drives upon exploratory behaviour. *J. comp. physiol. Psychol.*, **46**, 315–9.

MONTGOMERY, K. C. (1955). The relation between fear induced by novel stimulation and exploratory behaviour. *J. comp. physiol. Psychol.*, **48**, 254–60.

MONTGOMERY, K. C., and MONKMAN, J. A. (1955). The relation between fear and exploratory behaviour. *J. comp. physiol. Psychol.*, **48**, 132–6.

STEINBERG, H., RUSHTON, R., and TINSON, C. (1961). Modifications of the effects of amphitamene-barbiturate mixture by the past experience of rats. *Nature*, **192**, 533–5.

WATSON, R. J. (1964). Modification of the effects of drugs on behaviour by the nutritional state. In STEINBERG, H. (Ed.), *Animal Behaviour and Drug Action*. London: Churchill.

WILLIAMS, C. D., and KUCHTA, J. C. (1957). Exploratory behaviour in two mazes with dissimilar alternatives. *J. comp. physiol. Psychol.*, **50**, 509–13.

Manuscript received 6th March, 1967.

THE EFFECTS OF VARIATION OF INTERTRIAL INTERVAL ON EXPLORATION OF ELEVATED AND ENCLOSED MAZES

BY

M. S. HALLIDAY

From the Laboratory of Experimental Psychology, University of Sussex

The interval between exploratory trials was varied in experiments using simple elevated or enclosed mazes. Activity on the second trial was depressed for short inter-trial intervals but had recovered after about 10 min.; the degree of recovery was different in the two types of maze. A second decrease in second trial activity was found with inter-trial intervals of more than 20 min., but a further experiment suggested this was due to the effects of delay itself, rather than to previous experience of the environment.

Introduction

It has been shown that when rats are given daily exploratory trials in simple mazes there is a decline in activity from day to day in an enclosed maze (up to 3 days), but in an elevated maze exploratory activity remains constant (Halliday, 1967). The present experiments were designed to investigate this effect further by varying the intervals between trials in both elevated and enclosed mazes.

Experiment I

Subjects

The subjects were 133 naive female hooded rats, obtained from the Medical Research Council, Mill Hill, London. They were 100–120 days old at the time of the experiment. The rats were housed, four to a cage, in wire mesh cages, and had *ad lib* access to food (Laboratory chow) and water.

Apparatus

An elevated Y-maze was used with three arms meeting at angles of 120°, each arm was 18 in. long, 4 in. wide and 18 in. above floor level. The maze was placed under a metal framework, 4 ft. 6 in. × 4 ft. 6 in. × 4 ft. high, covered with black cloth. The subjects could be observed on the maze through a one-way screen 9 in. × 14 in. in the top of the enclosure. The enclosure was lighted by a single 100 watt bulb suspended centrally 6 in. from the top. The experimental room also contained a rack of solid metal cages with wire mesh fronts and floors. These cages measured 9 in. × 9 in. × 8½ in. high.

Masking noise of moderate intensity was supplied by the air conditioning equipment.

Procedure

The subjects were given 4 days pretraining during which they were allowed to move around in a 3 ft. square enclosure with metal walls, being picked up and handled from time to time.

On the fifth day the subjects were divided randomly into groups and all given two 3-min. trials on the Y-maze with a delay between trials. The groups differed in the length of the delay as shown in Table I. On each trial the rat was placed at the end of the same arm of the maze and allowed to explore for 3 min.

Each arm of the maze was divided, by a line on the floor, into two 9 in. sections and records of entries to these sections were made for each 30 sec. of the trial. At the end of 3 min., the rat was removed from the maze, put in one of the solid metal delay cages and left there for the length of time appropriate to its group. It was then given a second trial on the elevated maze and returned to its home cage. For delays of more than 5 min., food and water were available in the delay cages.

TABLE I

SUMMARY OF PROCEDURE. EXPERIMENT I

Group	*I*	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*
N =	10	10	21	12	23	12	24	10	12
Delay	¾ min.	2 min.	5 min.	10 min.	20 min.	40 min.	1 hr. 20 min.	5 hr. 20 min.	21 hr. 20 min.

The variation in numbers of subjects in each group was the result of the experiment being run in three parts:

Part A: Groups I, II and 10 subjects in Group III;

Part B: Eleven subjects in Group III, 12 in Group V, 12 in Group VII and all of Group VIII and IX;

Part C: Group IV, 11 subjects in Group V and 12 each in Group VI and VII.

Various parts of the experiment were run successively in order to extend the range and complete the picture which emerged from the earlier parts. But for the purposes of this report all parts have been treated together, since there was no difference in procedure.

All second trials were run between 2 p.m. and 4 p.m.; the first trials for all except Groups VIII and IX were run between 1.30 p.m. and 3.30 p.m. Group VIII had its first trial at about 10 a.m. and Group IX at about 5.45 p.m.

Results

The rats in Part B of the experiment, which had been obtained from the suppliers at a different time from those in Parts A and C, were significantly less active ($p < 0{\cdot}05$, U-test) on the first trial than those in the other two parts (which did not differ from one another). Presumably this difference was due either to genetic factors, or, more probably, to environmental differences before arrival at the laboratory. (All tests of significance are two-tailed unless otherwise stated.)

The first trial scores for the groups in Parts A and C combined and for the groups in Part B were each homogeneous (Kruskal-Wallis analysis of variance). There is thus no evidence for significant differences in exploratory activity arising from giving the first trials at different times of day; the daily cycles of activity commonly found in activity wheels (Wang, 1923) seem relatively unimportant in this type of situation. When a group was composed of subjects from Part B and Parts A or C, the changes in activity were in the same direction and of the same order of magnitude in both cases. It seems that, provided no direct comparisons of amounts of activity on the second trial are made between subjects in the differing parts of the experiment, valid comparisons between groups can be made. To avoid this difficulty tests of direction of change were used where appropriate.

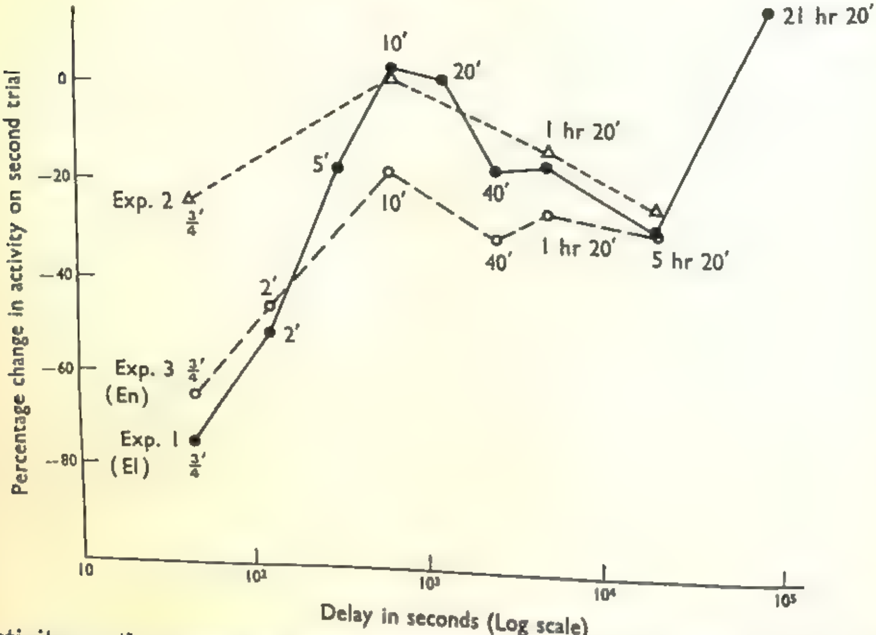
For similar reasons Figure 1 and Table II present the results in terms of percentage change in activity scores.

There was a sharp reduction in Trial 2 scores for short delay intervals; this reduction disappeared with delays of 10–20 min. There was also a reduction with delays between 40 min. and 5 hr. 20 min., but the 21 hr. 20 min. group showed no reduction. Separate Kruskal-Wallis analyses of variance were carried out on the Trial 2 scores for the groups in each part. In each case the variance was significant at below the 0·05 level.

The rise in Trial 2 scores from Group I to Group IV was significant; Group I (¾ min.) differed significantly from Group III (5 min.) ($p = 0{\cdot}02$, U-test). In

addition the following differences in direction of change of scores between the two trials were significant by χ^2 tests: I vs. IV ($p < 0{\cdot}01$), II vs. IV ($p < 0{\cdot}05$), II vs. V ($p < 0{\cdot}05$).

FIGURE 1



Activity on the second trial expressed as a percentage of activity on the first trial. Experiments I and II and III.

TABLE II

MEAN ACTIVITY SCORES AND PERCENTAGE CHANGE IN SCORES. EXPERIMENT I

Group	*Delay*	*Mean activity score*		*Percentage change in activity*	*p of change (signs test)*
		Trial 1	*Trial 2*		
I	¾ min.	18·6	4·6	−73·3	0·001
II	2 min.	18·2	8·8	−51·6	0·055
III	5 min.	16·7	13·76	−17·7	0·032
IV	10 min.	18·5	19·08	+3·2	—
V	20 min.	17·8	17·81	+0·25	—
VI	40 min.	19·0	15·42	−18·8	0·073
VII	1 hr. 20 min.	15·1	12·41	−17·7	0·058
VIII	5 hr. 20 min.	13·5	9·6	−28·9	—
IX	21 hr. 20 min.	10·0	11·7	+16·7	0·055

The fall in second trial scores after 20 min. delay was also significant. In this case, since the first trial scores for Groups VII and VIII were lower than those for Groups I–VI, it was inappropriate to use U-tests, but the χ^2 test for significance of changes shows that Group IV (10 min.) differed significantly from Group VI (40 min.) ($p = 0{\cdot}05$) and from Group VII (1 hr. 20 min.) ($p < 0{\cdot}01$). Furthermore there was a significant difference between Groups VI, VII and VIII combined and Group IV ($p < 0{\cdot}03$, χ^2) and also Group V ($p < 0{\cdot}05$, χ^2).

The changes in scores between trials for Group IX (21 hr. 20 min.) did not differ significantly by the χ^2 test from those in Groups VI, VII, and VIII individually, but they did when the latter three groups were combined ($p = 0{\cdot}02$, χ^2).

Discussion

With short intervals between trials in an elevated maze, there was a sharp decrease in exploratory activity on the second trial. With increasing delays this decrease became less and less apparent, until the delays of 10–20 min. between trials it disappeared altogether, and the rats were as active on the second as on the first trial. When the interval between trials was increased beyond about 20 min. activity on the second trial began to fall again up to delays of 5 hr. 20 min.; however with a delay of 21 hr. 20 min. the activity on the second trial was actually (non-significantly) more than on the first trial. These changes in activity do not appear to be attributable to any diurnal activity changes, since all the second trials were run at approximately the same time of day; while such differences in time as there necessarily were between times of first trials in the various groups gave no indication of producing differences in activity on that trial. The changes in activity on the second trial are obviously not easily accounted for by supposing that they were wholly due to changes in the novelty of the maze; nor indeed do they fit in with any explanation in terms of a single variable.

The most plausible account, leaving aside for the moment the 21 hr. 20 min. group, might be to say that the first part of the curve, up to 20 min. represents a recovery of exploration, due to renewal of novelty or for some other reason; while the subsequent depression of scores with delays up to 5 hr. 20 min. is the effect of being retained in the delay cages and is unrelated to the novelty or otherwise of the maze on Trial 2. The next experiment was performed to test this suggestion.

Experiment II

If the reduction in activity on Trial 2 with delays greater than 20 min. was an effect unrelated to the novelty of the stimuli during this trial, then it should make no difference, for these longer delays, if the two trials were to be given in different environments. On the other hand such an environmental difference should eliminate the decrease on the second trial with delays of less than 10 min. if this decrease is directly related to the stimulus characteristics of the environment. This prediction was tested by giving rats two trials with varying delays between them, the first trial in an open field and the second in an elevated Y-maze.

Subjects

The subjects were 48 naive female hooded rats, about 120 days old at the start of the experiment. They were housed and fed as in Experiment I.

Apparatus

The elevated Y-maze under the enclosure was as in Experiment I. An open field was used, 3 ft. × 3 ft. with 2 ft. high plain metal walls; the floor of the field was marked

out into nine 1 ft. × 1 ft. squares. This open field was placed on the floor of the experimental room and was illuminated by a single 150 wall bulb in the ceiling (about 11 ft. above the floor).

Procedure

This was in general the same as in the previous experiment. All subjects were given three days pretraining on a table top.

On the fourth day each rat was given two trials. Trial 1, lasting 3 min., was in the open field; this was followed after a period spent in the delay cages by a second trial in the Y-maze. In the open field each rat was started from the same corner; the rats' movements were marked on a cyclostyled sheet and the total number of line crossings made was recorded.

The subjects were randomly divided into four groups with different delay periods:

Group I	¾ min. delay	N = 12
Group II	10 min. delay	N = 12
Group III	40 min. delay	N = 12
Group IV	5 hr. 20 min. delay	N = 12

All second trials were given between 2 p.m. and 4 p.m.; the first trials were also given in the same period, except for subjects in Group IV which had their first trial at about 10.30 a.m.

Results

Table III gives the mean activity scores for each group on each trial.

TABLE III

Mean Activity Scores in Open Field and Y-Maze. Experiment II

Group	*Trial 1 Open field*	*Trial 2 Y-Maze*
I	25·33	14·33
II	25·1	18·9
III	25·6	16·4
IV	25·25	14·4

The scores for all groups on Trial 1 were very similar. On Trial 2 Group I differed from Group II and Group II from Group IV at the 0·1 level (U-test).

In order to use these results to help interpret those of the previous experiment it is necessary to find some standard for comparison. The 10 min. group in the last experiment gave virtually the same activity score on both trials (Trial 1 = 18·5; Trial 2 = 19·08); the 10 min. group (Group II) in this experiment also gave a Trial 2 score which was virtually the same (18·9). In order to make a comparison it was assumed that, had Group II (10 min.) in this experiment been given its first trial in the Y-maze, it would have scored about the same as it did on the second trial (i.e. about 19·0). It was also assumed that groups of rats which gave equal scores in the open field would, under identical conditions, give equal scores in the Y-maze. Both these assumptions are unproved but are reasonable. On this basis a "per cent. decrease" score was calculated for Trial 2 taking an assumed "Trial 1" score of 18·9 for all groups; these scores were: Group I—24·7 per cent.; Group II—No change; Group III—13·7 per cent.; Group IV—24·4 per cent. These results are shown along with the scores for the last experiment, in Figure 1.

The scores in this experiment from 10 min. onwards were virtually the same as those in the last experiment; the score for the $\frac{3}{4}$ min. group, however, though a little depressed on Trial 2, was much less so than the equivalent group in Experiment I. The difference on the second trial between the activity scores of the $\frac{3}{4}$ min. groups in the two experiments was highly significant ($p < 0.002$, U-test).

Discussion

These results confirm the proposed interpretation of the results of Experiment I. There seem to be two types of reduction in exploration with increasing delays, one of which, with delays of less than 10 min., only appears when the first trial is in the same environment as the second; the other, taking effect at delays of over 20 min., is a general effect of the delay procedure and appears equally when the first trial is in the same or a different environment to the second. This second decrease, for delays of between 20 min. and 5 hr. 20 min. is thus not a function of the stimuli in the environment, but rather of the delay itself. It is suggested that rats kept in the delay cages, which offer a rather monotonous environment, become les aroused as time goes on and that this decrease in arousal makes them less likely to be active in any exploratory situation, novel or familiar. An informal experiment in which rats' activity in a delay cage was measured, by means of a phototransistor and a beam of infra-red light, showed a steadily decreasing amount of activity in the 5 hr. after being put in the cage. This finding is consistent with the suggested interpretation. The recovery in activity with delays as long as 21 hr. 20 min. is puzzling; possibly, following the suggestions of Berlyne (1960) long periods of confinement in a dull environment produce boredom, which in turn leads to higher arousal.

While delays between trials of more than 10 min. produce decreases in activity on the second trial which are a function of the delay itself, delays shorter than this also produce a marked decrease when the rat has both trials in the Y-maze, but very little decrease when the first trial is in the open field. This decrement, in contrast to the decrements at longer delays, is dependent on the similarity between the environments on the two trials (cf. Montgomery, 1953). It has been shown (Halliday, 1967) that there is complete recovery of exploratory activity in an elevated maze between daily trials; this experiment shows that this recovery all takes place within about 10 min. after the end of the first trial. The implications of this finding will be discussed after reporting the next experiment.

Experiment III

While there is complete recovery in exploratory activity between daily trials in elevated mazes, there is a decrement in activity between such trials in an enclosed maze. This experiment was performed to show the course of recovery of exploratory activity in an enclosed maze for comparison with the results of Experiment I.

Subjects, apparatus and procedure

The subjects were 71 naive female hooded rats obtained from the Medical Research Council, Mill Hill, London. They were 100–120 days old at the beginning of the experiment.

An enclosed Y-maze was used with three 18 in. arms, 4 in. wide, meeting at angles of 120°. The walls of the maze were made of unpainted metal and were 6 in. high; the floor was wooden and the roof was made of transparent Perspex sections. The maze was placed 18 in. above the floor inside the same enclosure as in Experiment I. Other details of apparatus were similar to those in that experiment.

The procedure was identical to that in Experiment I. Each rat received two 3-min. trials in the enclosed maze; the numbers and delays in each group were as follows:

Group I	¾ min. delay	N = 12
Group II	5 min. delay	N = 11
Group III	10 min. delay	N = 12
Group IV	40 min. delay	N = 12
Group V	1 hr. 20 min. delay	N = 12
Group VI	5 hr. 20 min. delay	N = 12

Times of trials and recording procedures were as in Experiment I.

Results

The activity scores for both trials, the change in score on Trial 2 expressed as a percentage of the score on Trial 1 and the significance of the change are shown in Table IV.

TABLE IV

Mean Activity Scores and Percentage Changes in Scores. Experiment III

Group	*Delay*	*Mean activity scores*		*Percentage change in activity*	p *of change (Signs test)*
		Trial 1	*Trial 2*		
I	¾ min.	31·2	10·8	−65·6	<0·001
II	5 min.	24·3	13·7	−45·7	0·021
III	10 min.	27·2	22·4	−17·6	0·038
IV	40 min.	26·75	18·4	−32·0	0·012
V	1 hr. 20 min.	27·9	20·8	−25·5	0·05
VI	5 hr. 20 min.	28·1	20·2	−28·2	<0·01

A Kruskal-Wallis analysis of variance showed no significant variability between groups of Trial 1; this suggests that no important differences in exploratory activity arise from the different times of day at which Trial 1 was necessarily given to the different groups.

A Kruskal-Wallis analysis of variance of scores on Trial 2 just failed to be significant at the 0·1 level; this is not surprising since any variability must have arisen largely from Groups I and II only. The difference between Trial 2 scores for Group I and II was not significant; the significances of the differences for Group I vs. Group III and Group II vs. Group III were $p = 0{\cdot}02$ and $p = 0{\cdot}05$ (one-tailed, U-test). None of the differences in Trial 2 scores between Group III and Group IV, V and VI were significant; although all three of these groups were less active on Trial 2 than Group III, both in terms of absolute scores and percentage change in score.

Discussion

Since the procedure in this experiment was identical to that in Experiment I, it is appropriate to make a fairly close comparison of the results. Figure 1 plots the results of both experiments expressed as percentage change in score on Trial 2 for the various groups. There is clearly a close similarity in the pattern of results. There can be little doubt that the reduction in Trial 2 scores with delays of longer than 20 min. is a similar effect to that found in the two previous experiments; it is

therefore not of direct interest for present purposes and will not be considered further. In the elevated maze it was found that the decreases in Trial 2 scores at intervals shorter than 10 min. were a function of the stimuli in the situation (i.e. they did not appear when the two trials were given in different environments). The same is presumably true in this case. For the analysis of exploratory behaviour in these two situations attention should therefore be concentrated on the groups with delays of 10 min. or less.

In the elevated maze there was a complete recovery in exploratory activity after a delay of 10 min., but in the enclosed maze there was only about 80 per cent. recovery after the same delay. A significantly greater proportion of animals in Group III (10 min.) of Experiment III reduced their scores between trials than did in the 10-min. delay group in the elevated maze ($p < 0{\cdot}05$, Fisher Test). With daily trials there is no intertrial reduction in activity in elevated mazes, while in enclosed mazes the percentage reduction between the first 2 days is approximately the same as after only 10 min. in this experiment (Halliday, 1967). Thus in both types of maze the effects of a 10-min. delay are more or less equivalent to the effects of a day's interval between trials. It is concluded that recovery of exploratory behaviour, after a 3-min. trial at least, all takes place within about 10 min.

One explanation of these results is that the changes in this brief 10-min. period arise because the rats forget most of what they learnt during the first trial, and so the maze is more or less novel to them after 10 min. This seems to be an unreasonably fast rate of forgetting and makes it hard to account for any sort of learning in an elevated maze; for, unless information acquired during exploration and information acquired by reinforcement are *stored* in different ways, the rate of forgetting should be the same. There are other difficulties about such an account. The percentage reduction in Trial 2 scores (as compared to Trial 1 scores) is very similar with $\frac{3}{4}$ and 2 min. delays in both elevated and enclosed mazes. If differences in exploratory activity in the same maze are taken to reflect differences in the novelty or familiarity of the maze, it follows that, regardless of any initial differences in "interestingness" between the mazes, the relative decrease in novelty (increase in familiarity) $\frac{3}{4}$ min. after a 3-min. trial in either maze is about the same. That is the novelty of the environment $\frac{3}{4}$ min. after a 3-min. trial, compared to its initial novelty at the beginning of the first trial, is the same in both types of maze. The question posed by these results is: if the relative novelty of the two types of maze is about the same $\frac{3}{4}$ min. after a 3-min. trial, why is it different 10 min. afterwards? It appears that, unless very implausible assumptions are made about differential rates of forgetting in the two types of maze, it is very hard to account for these results simply in terms of what the animals forget in 10 min. Such a conclusion throws serious doubt on any theory of exploration in rats which assumes that what the rat knows about the maze is the sole determinant of exploratory behaviour.

This work was carried out as part of a programme supported by the Medical Council. I am grateful to Mr. A. J. Watson for his help and advice.

References

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw-Hill.

Halliday, M. S. (1967). Exploratory behaviour in elevated and enclosed mazes. *Quart J. exp. Psychol.*, **19**, 254–63.

Montgomery, K. C. (1953). Exploratory behaviour as a function of 'similarity' of stimulus situations. *J. comp. physiol. Psychol.*, **46**, 315–9.

Wang, G. H. (1923). Relation between 'spontaneous activity' and oestrus cycle in the white rat. *Comp. Psychol. Monogr.*, **2**, 6.

Manuscript received 6th March, 1967.

THE TIME COURSE OF PREPARATION*

BY

PAUL BERTELSON

From the M.R.C. Applied Psychology Research Unit, Cambridge†

The time course of the adjustments triggered by a warning signal was studied by measuring choice reaction times (RTs) at different predictable foreperiods after such a signal. Before the warning signal, a high time uncertainty situation was created by imposing either a long constant foreperiod of 5 sec. or one varying in the range 1·5 to 5 sec. The warning signal was a click. Foreperiods ranging from 0 to 300 millisec. were used in different blocks of trials. The stimulus was the onset of one of two lamps calling for the pressing of one of two keys. A control condition, without click, was used also. RTs were found to decrease continuously when the foreperiod was increased from 0 to 100–150 millisec. The click delivered simultaneously with the stimulus permitted reactions significantly faster than in the control condition. It is concluded (*a*) that the latency of preparation can be much shorter than the 2 to 4 sec. reported by Woodrow; (*b*) that the warning signal can be used as a time cue to start preparatory adjustments without starting a refractory period of the order of magnitude found in experiments with pairs of successive reactions, and thus that such refractory periods are not the inevitable cost of paying attention to a signal. There is also some suggestion that in this situation the click not only triggers preparatory adjustments, but also causes an immediate facilitation of the reaction to the visual stimulus.

INTRODUCTION

Reaction time (RT) is a function of uncertainty about the time of occurrence of the stimulus, or time uncertainty. The effect has been studied for both the simple (Botwinick and Brinley, 1962; Kay and Weiss, 1961; Karlin, 1959; Klemmer, 1956; Woodrow, 1914) and the choice reaction (Bertelson and Boons, 1960; Bertelson and Borsu, 1965; Boons and Bertelson, 1961; Broadbent and Gregory, 1965) by manipulating the *foreperiod* intervening between a *warning signal* and the stimulus. It can be interpreted as showing that the warning signal is used as a time cue to start some preparatory adjustments, either immediately or shortly before the expected time of occurrence of the stimulus.

A question which has been rather neglected is how long it takes for these adjustments to build up. Woodrow (1914) obtained a slightly longer RT with a constant foreperiod of 1 sec. than with one of 2 sec. for each of his three subjects. For one subject, the shortest RT was obtained at 4 sec. Woodrow concluded that it takes between 2 and 4 sec. to reach full "attention." The existence of an optimum of about that duration for the foreperiod has been accepted by most subsequent reviewers (Chocholle, 1963; Teichner, 1954; Woodworth and Schlosberg, 1954), but it is contradicted by more recent work: for simple reactions, 1 sec. is better than 2 sec. (Klemmer, 1956), 0·5 sec. is better than 1 sec. (Botwinick and Brinley, 1962; Karlin, 1959; Sanders, 1965); for choice reactions, 0·5 sec. is better than 2·5 sec. (Boons and Bertelson, 1961). Klemmer considered the possibility that Woodrow's finding might have resulted from the fact that the warning signal itself was given at irregular intervals, but an attempt to produce the effect experimentally failed. An alternative

* The present experiment has been described at the Meeting of the Experimental Psychology Society held in Oxford in July, 1965, and at the XVIIIth International Congress of Psychology in Moscow, in August, 1966.

† On visit from the University of Brussels. Now returned to the Psychological Laboratory of this University.

explanation for Woodrow's results would be that the ascending-descending order of presentation of the different foreperiods within a session put the 1 sec. foreperiod, the shortest one, in an unfavourable position.

Recent work thus shows that preparation can be built up much faster than Woodrow contended, but it has produced no comprehensive picture of the time course of the phenomenon, for there has been no systematic study of the effects of the duration of constant foreperiods in the range below 1 sec. Shorter foreperiods have only been studied in experiments with variable foreperiods. Lansing, Schwartz and Lindsley (1959) measured simple RTs to flashes preceded at a variable interval in the range 50 to 1,000 millisec. by another flash, and found an optimum at about 350 millisec. The longer RTs at shorter intervals might be due to lack of time to prepare, but also to expectancy regarding the time of occurrence of the stimulus. The last possibility is demonstrated by an experiment by Drazin (1961) who measured simple RT to a light preceded by a warning tone, the range of variation of the foreperiod covering 1 sec. in all conditions, but starting at values extending from 0·125 to 2 sec. All reaction time-foreperiod curves displayed a marked initial gradient. However, the gradient was somewhat steeper when the shortest foreperiod was less than 250 millisec. which suggests that a latency phenomenon is superimposed on the expectancy effect. To study the time course of preparation, uncontaminated by range effects, one must of course use constant foreperiods. One reason why it has not been done is that most authors studied simple reactions. In that situation, with short constant foreperiods, it is difficult to prevent the subject from reacting to the warning signal. The obvious way to avoid this type of behaviour is to attach event uncertainty to the stimulus, i.e. to use choice reactions.

In the experiment to be reported, a high time uncertainty was first created, by imposing a long waiting delay, or a highly variable one. A visual stimulus calling for a choice reaction was then presented, either without warning or a constant foreperiod after an auditory warning signal. By varying the length of the foreperiod, it was thus possible to observe the time course of the adjustments triggered by the warning signal.

An additional motive to study this situation was provided by the consideration of the current work on the *psychological refractory period*. It has been conclusively shown that when two signals, each calling for a response, occur at an interval shorter, or sometimes even slightly longer, than the RT to the first signal, the response to the second signal is delayed, provided the subject has been persuaded to deal immediately with the first signal. This is the main experimental basis for the hypothesis, first put forth by Craik, that information processing involves a basically intermittent central mechanism (for a recent review, see Bertelson, 1966).

"Intermittent" means that once one of the unitary operations performed by the central mechanism has been started, it cannot be stopped, and further sensory input cannot be accepted, until its full completion. In the case of overt immediate reactions, the intermittency period would consist of the main part of the RT, plus possibly some recovery time (Davis, 1957). It has been suggested that other operations performed on a signal, like eliminating irrelevant stimuli, or storing in short-term memory, might involve intermittency periods of similar duration. Davis (1959), for example, has argued that paying attention, not responding, is critical in determining refractory delays. He based this conclusion on experiments, by Fraisse (1957) and by himself, where delays were found in second RTs, not only in the usual situation where both signals are responded to, but also when the second signal only called for a response. Davis did not contend that any irrelevant signal would disturb subsequent reactions, and he discussed the reasons why such disturb-

ance was found in his experiment. In fact, there are cases in the literature where irrelevant signals caused no refractoriness (Borger, 1963; Rubinstein, 1964). But implicit in Davis's discussion is the idea that signals either engage the channel for a whole intermittency period, or not at all. At first sight, this idea seems inconsistent with the fact that mean delays vary from situation to situation: they are longer in the two-responses case than in the second-response case (Fraisse, 1957); they are longer when both stimuli reach the same sense modality than when they reach different modalities (Fraisse, 1957; Davis, 1959). Such variations can however still be accounted for by combinations in varying proportions of cases of full refractoriness and of cases of no refractoriness.

When the case of warning signals is considered in the framework of this discussion, one is immediately led to ask if a signal can be used as a time cue to start preparation without causing refractoriness. More precisely, does the warning signal start an intermittent period, so that the processing of the stimulus itself must be delayed till preparation is completed? If it does, at short foreperiods, like 50 millisec., the gain from preparation might be smaller than its cost. The subject would then have the choice between ignoring the warning signal, and having the same long RT as without warning, or accepting it at the cost of a still longer RT. So, the warning signal would either be ineffective, or harmful, below a certain length of foreperiod.

Method

Apparatus

Two neon pea-lamps were mounted 3 cm. apart on one horizontal axis, at about eye level, behind a vertical ground-glass screen. They provided the two stimuli. Their position was indicated on the screen by outline circles. The response unit consisted of two horizontal keys, 2·5 cm. apart, on which the subject rested the index and middle fingers of his preferred hand. The correct response was to press the right key for the right lamp and the left key for the left lamp. The subject wore padded earphones, through which 85 Db. white noise was applied, to mask all noises from the programming equipment, which was situated in the same room. The click, which served as warning signal, was superimposed on the white noise.

A trial started when the experimenter activated a switch, which started the white noise and cancelled the stimulus light of the preceding trial. After a *waiting delay*, the click was applied (in the *experimental trials*) and was followed, after the *foreperiod*, by the stimulus. This stayed on until the experimenter started the following trial. In the *control trials*, no click was applied, and the stimulus came immediately at the end of the waiting delay.

The waiting delays and the foreperiods were controlled by two Labgear dekatron interval timers. RTs were measured on a Labgear dekatron timer. The sequence of stimuli was controlled by a Creed punch tape reader.

Subjects

Ten Royal Naval ratings participated in seven sessions each, normally one in the morning and one in the afternoon of 3½ consecutive days.

Procedure

Measurements were made under two waiting delay conditions: *constant* waiting delays of 5 sec. or *variable* waiting delays; in the latter condition, the eight delays ranging by 0·5 sec. steps from 1·5 to 5·0 sec. were used with roughly equal frequencies, in random order. Each subject had three experimental sessions under each condition, in alternated order; five subjects started with the constant condition and five with the variable one.

At each experimental session, i.e. sessions 2 to 7, seven foreperiod durations were used: 0 (i.e. simultaneous presentation of the click and of the stimulus), 20, 50, 100, 150, 200 and 300 millisec. One block of trials was run with each foreperiod, together with two blocks of control trials. One block consisted in principle of 12 trials, of which the two

first were considered as examples enabling the subject to experience the foreperiod. If the subject made errors in the remaining 10 trials, additional trials were given, to obtain a total of 10 errorless trials.

After each response, the experimenter stopped the masking noise and told the subject whether his reaction time had been "fast" or "slow," meaning faster or slower than the mean correct RT with the same foreperiod on the previous session. When an error was made, the RT was not communicated.

The order of the blocks within sessions was counter-balanced, using two 9 × 9 latin squares, as far as the number of sessions permitted.

The first session was devoted to familiarization with the task. The procedure was varied from subject to subject.

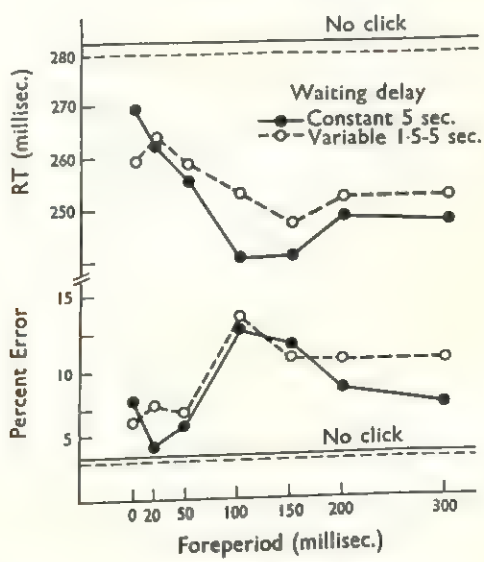
The subjects were tested in two groups of respectively six and four. In each group, individual mean RTs for the preceding sessions were displayed in diagrammatical form on the door of the experimental room before the start of the next session.

Results

Effect of warning click and of foreperiod duration

Figure 1 shows mean correct RTs, and percentages errors, averaged over subjects and sessions, as a function of the foreperiod duration, together with the corresponding data for the control (no-click) trials. The first two trials of a block were always discarded. Mean correct RTs were obtained from the following 10 correct reactions. Error rates were calculated from the 10 first reactions following the two practice trials, and are thus not influenced by the additional trials given to obtain the total of 10 correct reactions.

FIGURE 1



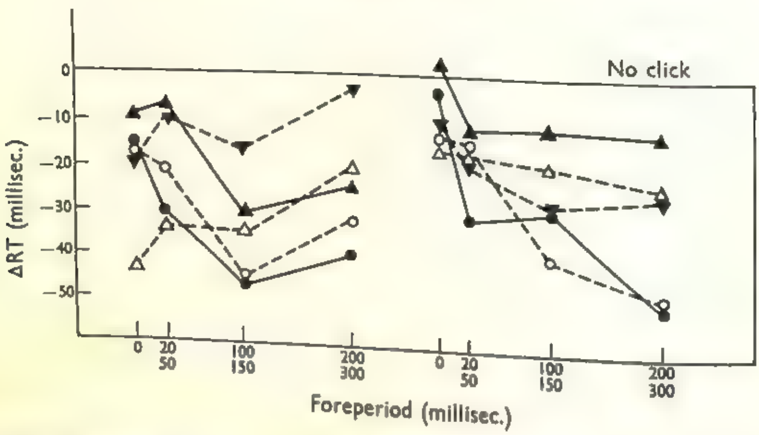
Mean correct reaction times and error percentages as functions of foreperiod duration, with waiting delay condition as parameter. The horizontal lines give the values for the control trials.

Under both waiting delay conditions, the RT decreases with increasing foreperiods, up to 100–150 millisec., then increases slightly again. On the other hand, it is shorter at foreperiod = 0 than on the control trials, where the timing was the same,

but no click was applied. Reductions in RTs are paralleled by marked increases in errors rates, in much the same fashion under both waiting delay conditions.

Large individual differences in the pattern of RTs were observed, as is apparent in Figure 2, which gives the RTs for the experimental trials as deviations from the values for the control trials. For this figure, the data from the two waiting delay conditions have been pooled.

FIGURE 2



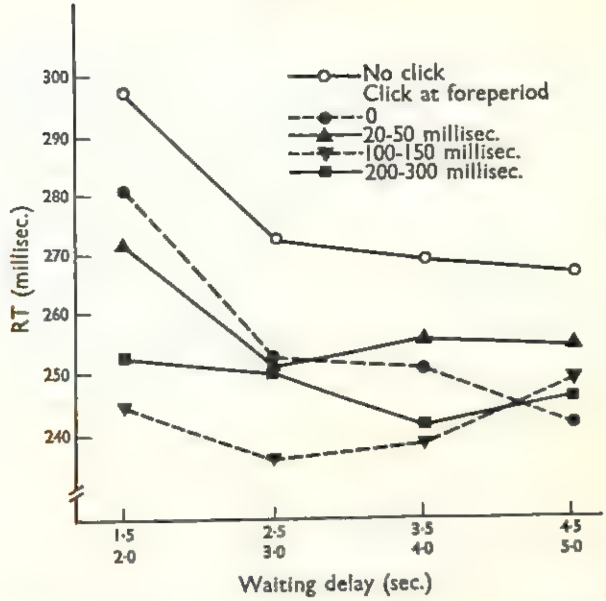
Mean correct reaction times on the experimental trials as functions of foreperiod duration. Results of individual subjects. The RTs are expressed (in millisec.) as deviations from the mean RT of the subject on the control trials.

Analysis of variance was applied to the mean correct reaction times per subject, waiting delay and warning condition (there are eight warning conditions: the seven foreperiods and the control condition). The interaction between waiting delay condition and warning condition is non-significant. There is a slight but significant ($p = 0{\cdot}05$) interaction between warning condition and subjects. When tested against this interaction, warning conditions still give a highly significant F (14·2 with 7 and 63 *d.f.*, $p = 0{\cdot}01$). Subsequent application of Duncan's multiple range test shows that, at $p = 0{\cdot}05$, (*a*) RTs in the control trials are longer than at any foreperiod duration; (*b*) RTs at foreperiod = 0 and 20, are larger than at foreperiods 100, 150, 200 and 300; (*c*) RTs at foreperiod = 50 are longer than at foreperiods 100 and 150. Other differences fall short of the 0·05 level.

Effect of the length of the waiting delay, in the variable condition

The data for the variable waiting delay condition were analysed as a function of the length of the waiting delay and of the warning condition. As is shown in Figure 3, without warning click and also with the simultaneous click, the RT decreases with the length of the waiting delay. A similar gradient has been reported by Breitwieser (1911), Karlin (1959), Drazin (1961) and Audley, Pike and Stillitz (1965). With the warning click preceding at increasing foreperiods, the effect of the waiting delay becomes progressively smaller. It is practically absent with foreperiods longer than 100 millisec. No effect of the waiting delay on the percentage errors was found in any warning condition.

FIGURE 3



Mean correct reaction times in the variable waiting delay condition, as functions of waiting delay duration, with warning condition as parameter.

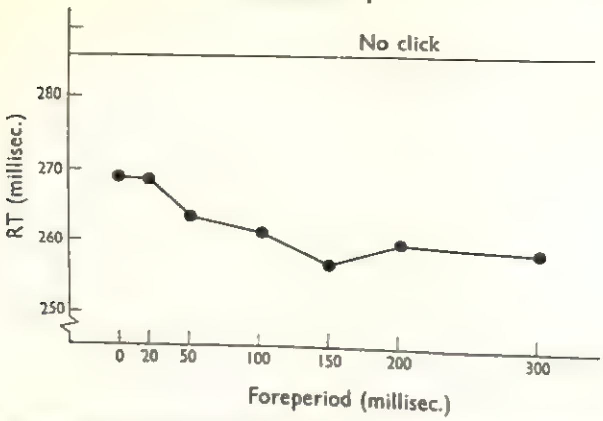
Discussion

The present results confirm the suggestion, derived from recent work on time uncertainty, that the latency of preparation can be much shorter than the 2 to 4 sec. value reported by Woodrow. The full effect of the warning signal on the reaction time was obtained after between 100 and 150 millisec. Whether this figure would hold for other reactions than the one which has been studied here—e.g. for higher degrees of choice or other S–R codes—cannot be stated.

In this experiment, the effect of the warning signal was not only to reduce the RT but also to increase the percentage of errors. Inspection of RT distributions showed that errors are generally accompanied by very short values. One might argue that the increase in speed was obtained by giving simple reactions to the onset of the light, thus disregarding the nature of the stimulus, on an increasing proportion of trials. If this crude hypothesis were true, a correction for guessing could be applied to the data. The principle was, for each error, to discard one correct reaction given with the same RT. In practice, the sum of correct RTs was subtracted from the sum of incorrect RTs, and means calculated on these corrected sums, separately for each block of trials. The results of these treatments are shown in Figure 4. It appears that the effect of the warning signal cannot be "explained away" by the hypothesis of random responding.

According to the reasoning developed in the Introduction, the fact that very short foreperiods like 50 millisec. and less already produce a detectable effect on reaction time implies that the warning signal can be accepted as a time cue without creating refractoriness. Let us suppose that processing the warning signal started an intermittent period of the order of magnitude obtained in experiments with pairs of simple reactions, say 150 millisec. Let us make allowance for the fact that auditory data reach the brain faster than visual ones, and let us take as an estimate of this difference the usual difference between simple visual and auditory RTs, 40 millisec.

FIGURE 4



Mean reaction times after "correction for guessing," as functions of foreperiod duration.

A foreperiod of x millisec. would result in an interval of $x + 40$ millisec. at the central level. If the gain from preparation is 40 millisec. (the mean value observed in this experiment) and if preparation occupies the channel for 150 millisec., the cost cancels the gain for all foreperiods shorter than 70 millisec.

A striking result is the effect of a click delivered simultaneously with the stimulus. The simplest explanation for this finding is in terms of the difference in conduction times to the cortex between visual and auditory stimuli, which were mentioned in the previous paragraph. The simultaneous click would thus produce a genuine warning effect at the central level. But it is difficult to explain the whole of the observed effect in this way. The mean reduction in RT between foreperiods 0 and 100 millisec. is about 17 millisec. The difference between the control condition and foreperiod 0 is also 17 millisec., thus much more than one would predict by extrapolating the apparently linear RT–FP curve up to a −40 millisec. FP, corresponding to the hypothesized central asynchrony. Of course, the reasoning which derives this asynchrony from the difference between the corresponding simple RTs is controversial, but, if anything, the estimate is rather generous. The data on subjective simultaneity of a click and a flash (Hirsch and Fraisse, 1964) would lead to smaller estimations.

One is then led to consider the possibility that, beside its function as a time cue, used to start preparation, the click also produces some immediate facilitation of the reaction to the visual stimulus. Hershenson (1962) has shown that a simple reaction to a visual stimulus can be facilitated by an auditory stimulus, the effect being maximal when the interval between the stimuli was equal to the difference between the RTs to the two stimuli. The effect was of very short duration, for no facilitation was detected when the click was simultaneous with the visual stimulus. Whether an effect of longer duration operates in the conditions of the present experiment remains to be demonstrated by further experiments.

Some support for the facilitation hypothesis can be derived from the interaction between waiting delays and warning conditions (Fig. 3). As was noted before, the click preceding the stimulus at foreperiods equal or larger than 100 millisec. had a much larger effect on the reaction time when the waiting delay was short, as if it cancelled the effect of the length of the waiting delay. This seems to mean that when a sufficient foreperiod is available, the expectancy regarding the length of the waiting delay ceases to play its usual role. The simultaneous click on the other

hand works independently of the length of the waiting delay, as if its effect was simply superimposed on that of expectancy.

This work has been partly supported by the Belgian Fonds National de la Recherche Scientifique, under the form of a "Crédit aux Chercheurs." Thanks are due to the Medical Research Council for accepting the author as a visiting member of this Unit. The visit was facilitated by a NATO research fellowship. The author is indebted to Sir Frederic Bartlett and to Dr. D. E. Broadbent for providing space, facilities, suggestions and encouragement. The work benefited much from discussions with members of the Unit.

References

Audley, R. J., Pike, A. R., and Stilitz, I. B. (1965). Effect of foreperiod distribution on simple reaction time. Paper read at the March, 1965 meeting of the *Experimental Psychology Society* in Reading.

Bertelson, P., (1966). Central intermittency twenty years later. *Quart. J. exp. Psychol.*, **18**, 153–63.

Bertelson, P., and Boons, J.-P. (1960). Time uncertainty and choice reaction time. *Nature*, **187**, 531–2.

Bertelson, P., and Borsu, M. (1965). Incertitude temporelle et reconnaissance tachistoscopique. *Année psychol.*, **65**, 17–26.

Boons, J.-P., and Bertelson, P. (1961). L'influence de l'incertitude temporelle sur le temps de réaction de choix. *Année psychol.*, **61**, 361–76.

Borger, R. (1963). The refractory period and serial choice reactions. *Quart. J. exp. Psychol.*, **15**, 1–12.

Botwinick, J., and Brinley, J. F. (1962). An analysis of set in relation to reaction time. *J. exp. Psychol.*, **63**, 568–74.

Breitwieser, J. V. (1911). Attention and movement in reaction time. *Arch. Psychol.*, **18**.

Broadbent, D. E., and Gregory, M. (1965). On the interaction of S–R compatibility with other variables affecting reaction time. *Brit. J. Psychol.*, **56**, 61–7.

Chocholle, R. (1963). Les temps de réaction. In Fraisse, P., and Piaget, J. (Ed.) *Traité de Psychologie expérimentale*, vol. 2, Paris: Presses universitaires de France, p. 81.

Davis, R. (1957). The human operator as a single channel information system. *Quart. J. exp. Psychol.*, **9**, 119–29.

Davis, R. (1959). The role of "attention" in the psychological refractory period. *Quart. J. exp. Psychol.*, **11**, 211–20.

Drazin, D. H. (1961). Effects of foreperiod, foreperiod variability and probability of stimulus occurrence on simple reaction time. *J. exp. Psychol.*, **62**, 43–50.

Fraisse, P. (1957). La période réfractaire psychologique. *Année psychol.*, **57**, 315–28.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *J. exp. Psychol.*, **63**, 289–93.

Hirsch, I. J., and Fraisse, P., (1964). Simultanéité et succession de stimuli hétérogènes. *Année psychol.*, **64**, 1–19.

Karlin, L. (1959). Reaction time as a function of foreperiod duration and variability. *J. exp. Psychol.*, **58**, 185–91.

Kay, H., and Weiss, A. D. (1961). Relationship between simple and serial reaction time. *Nature*, **191**, 790–1.

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *J. exp. Psychol.*, **51**, 179–84.

Lansing, R. W., Schwartz, E., and Lindsley, D. B. (1959). Reaction time and EEG activation under alerted and non alerted conditions. *J. exp. Psychol.*, **58**, 1–7.

Rubinstein, L. (1964). Intersensory and intrasensory effects in simple reaction time. *Percept. mot. Skills*, **18**, 159–72.

Sanders, A. F. (1965). Prewarning signal activity and RT as a function of foreperiod. *Percept. Mot. Skills*, **21**, 405–6.

Teichner, W. H., (1954). Recent studies of simple reaction time. *Psychol. Bull.*, **51**, 128–49.

Woodworth, R. S., and Schlosberg, H. (1954). *Experimental Psychology*. New York: Holt, p. 29.

Woodrow, H. (1914). The measurement of attention. *Psychol. Monogr.*, **17** (5, whole No. 76).

Manuscript received 20th March, 1967

BOOK REVIEWS

Nervous and Hormonal Mechanisms of Integration: Symposia of the Society for Experimental Biology, No. 20. Edited by G. M. Hughes. London: Cambridge University Press. 1966. Pp. viii + 553. 75s.

In 1950 the Society for Experimental Biology held a symposium on physiological mechanisms in animal behaviour. This made a great impression both on zoologists and psychologists. The reviewer attended while a research student, and it is difficult to exaggerate the feeling of excitement and stimulation which the conference aroused. The papers were organized to cover sensory mechanisms, sensory-motor integration, instinctive behaviour, and neural mechanisms and learning. Contributions by Lorenz, Tinbergen and Baerends gave perhaps the first taste of ethology to many English workers. Since then this discipline has become perhaps the major formative influence in behaviour studies. However learning was not neglected, being covered by papers from Thorpe, Konorski, Lashley and J. Z. Young. The symposium produced a genuine integration, indicating the basic problems in behaviour mechanisms, suggesting methods which could be useful in studying them, and providing the broad conceptual framework of ethology into which they fitted.

After 1950 there was an explosive development of work on behavioural mechanisms. This is well illustrated by the fact that Young's early paper on brain mechanisms and learning in octopuses quoted only one similar reference, while Wells's competent review in 1966 quotes at least 53! Although most of the subsequent octopus work employs methods which were available 17 years ago, progress in other fields has been very much influenced by the development of techniques such as electron microscopy, and particularly electrophysiological recording. It is significant that of the 19 papers in this symposium, 15 report electrophysiological findings. The symposium is also biased towards invertebrates, only four of the papers being concerned with vertebrates, and more than half the others involving arthropods. This bias is understandable since invertebrates are technically easy to work with, and preparations can be made which involve only a few simple neuronal circuits, suitable for experiments on neuronal interaction and integration.

The field of invertebrate reflexology studied both behaviourally and electrophysiologically is well represented here in papers by Kennedy *et al.*, Wilson, Wendler, and Josephson. Maynard describes a detailed analysis of the integrative properties of neuronal circuits composed of several cells, in a paper important not only for the elegance of the analysis itself, but for revealing several parameters of electrical activity which affect synaptic excitability. Horridge describes recent advances in his study of the optomotor responses of the eyestalk of crabs, showing how control theory can be fruitfully applied to this system. His results suggest that the original type of model for movement perception based on the work of the Reichardt–Hassenstein approach may prove inadequate. He has also shown the existence of a short term memory for movement, which raises intriguing possibilities. Electron microscopy was used to study the synaptology of the peripheral visual system in crabs, and it is clear that this when combined with the electrophysiological analysis, which is currently under way, will prove very fruitful.

The sensory side of behaviour is covered by three papers. Schneider deals with chemoreception in insects, and is particularly concerned with the receptor responses to chemicals which are biologically important in communication. Wiersma describes the properties of the crustacean visual pathways; it is interesting to compare the coding of information in crabs and lobsters with that of vertebrates. Roeder and Payne have extended their work on "bat detection by moths" in an admirable way, combining a study of the structure and function of the auditory mechanisms with its detailed physiological working, and paying due regard to its adaptive consequences. This provides a true synthesis of various methodological and conceptual approaches.

The coverage of more complex areas of integration, relating behaviour to neural and hormonal mechanisms is less satisfactory. There is one paper on chemical transmitters on invertebrates (a field which needed reviewing) and another on the secretion of hormones by neurones. However, the only two other papers in this area are reviews of hormonal factors in the temporal integration of reproductive behaviour in canaries (by Hinde and Steel) and sticklebacks (Baggermann). Both of these use sophisticated ethological measures, and show what can be achieved in this area. Unfortunately no obvious link is provided between the behavioural and the central nervous mechanisms, although work

in this area (for example by Sawyer, Harris and many others), has made significant advances. Only the paper by Loher and Hüber on reproductive behaviour in a grasshopper begins to achieve this sort of synthesis, while Clark provides a competent survey of the integrative action of the worm's brain.

The papers on learning are unrepresentative of their field. Wells' summary of the octopus work is useful. But vertebrate learning is represented by a study of electrophysiological plasticity in the spinal cat (Kozak and Westerman). Bruner and Tauc present a careful study of long lasting post-synaptic changes in Aplysia, and couple this with some simple experiments on learning. This preparation provides promising material for the analysis of the electrical and biochemical changes correlated with learning. However, after the inclusion of Lashley's contribution in the 1950 symposium it seems astonishing that the present volume contains no paper on the neurophysiological basis of learning in mammals—no mention of studies on the frontal, temporal or hippocampal lobes, the chemistry of learning, or the effect of brain stimulation on reinforcement. The omission seems extraordinary.

It is perhaps surprising that after reading the papers in this symposium, each one of which is good in its own right, the prevailing feeling is one of disappointment. After 1950 we felt that an integrated approach to the mechanisms of behaviour was indeed possible and this is reflected in Bullock's exciting introduction. Yet in the present symposium perhaps five of the 19 papers go some way to adopting this approach, and the work of physiologists and psychologists which could have bridged some of the gaps was not included. The despondency is well summarized by the editor in his preface where he says: "This meeting has served to emphasize the need to apply the full gamut of neuro-endocrinological, electrophysiological, neuroanatomical, ethological and cybernetic techniques during future work in this active field. The need for an integrated approach to integration was amply demonstrated." Psychologists may well feel sad that their own science is omitted from the editor's list. Perhaps zoologists should be reminded that there are (at a guess) some 10 good journals in the broad field of physiological psychology. It would be sad if the closing gap between zoology and psychology should now be widened.

D. M. Vowles.

Essential Works of Pavlov. Edited with an introduction by Michael Kaplan. New York: Bantam Books. 1966. Pp. xiii + 392. $1.25.

A paperback edition of selected Pavlov papers is welcome, especially when the selection made can provide a rational summary of the work and ideas. The writings chosen here give a lucid account of Pavlov's work as it developed from the irregularities which he found in his earlier work on digestion: as we all know, he found he could deal quantitatively with the relation between physical stimulation and the secretion of saliva, but there was no way of dealing with mouths watering at the sound of the experimenter's footsteps except in subjective, psychological terms—thoughts, feelings and desires—which he rejected as scarcely amenable to scientific analysis. He had learned to consider an animal's activity physiologically, as the activity of the nervous system, and found he could not abandon this position when faced with the wider problems of what is still frequently regarded as behaviour *as distinct* from physiology. In enviably concrete and lucid terms he discusses the difficulties in dealing with these increasingly precise and delicate forms of adaptation, and the consequent challenge (more clearly marked in late nineteenth-century Russia than we know it today?) involved in extending a rigorously scientific approach in an area confused by the traditional tendency to think separately of mind and body.

The selection reprinted here outlines the clear, several sided idea central in Pavlov's work on conditioning: that the brain is the organ for maintaining the equilibrium between an organism and its environment which is necessary for survival; that behaviour is higher nervous activity, reflex and therefore determined largely by the environment; that the idea of physiological reflexes can be extended in different ways to provide for delicate adjustments to an infinite variety of stimulation, and for the cumulative effects of previous adaptation or experience. With the ideas goes the technique: the activity of rather unimportant, therefore normally rather isolated reflexes like the salivary ones can by rigorous systematization be made into a tool both to explore, on the basis of what is known of neural anatomy, the activity of the central nervous system, *and* to bring into some order the changing active relationship between an organism and different aspects of its environment which is behaviour. This, the development of ideas and method, the

organizing of results, the struggle with psychologists who persist under one guise or another in their old, dualistic ways, and the excursion into psychiatry cover the remainder of Pavlov's part in this book.

We learn that this edition was undertaken because the editor, working in the field of American conditioning referred to the original papers, found them richer scientifically than he expected, and realized that allusions to Pavlov are frequently misinformed. Unfortunately the editorial material presented here, although clearly earnest in intention does little to orient the reader towards a genuine understanding of Pavlov's position. The introductory chapters beside some biographical material and excursions on to subjects like teaching machines and the Korean war attempts to draw up a general scheme in psychology where Pavlov, Freud, and B. F. Skinner make parallel and complementary contributions. Apart from that Pavlov has been left to speak for himself which is, of course, good. But Pavlov has been speaking for himself and (at any rate, in the Anrep translation) in beautiful lucid English for between 40 and 50 years. People just do not seem to grasp what he was saying. He regarded himself as an innovator, with an approach wholly at variance with and scientifically more correct than that of most psychologists. At some stage a full analysis of his contribution which shows up the differences as well as points of contact would be more than welcome. M. VINCE.

Psycholinguistic Papers. Proceedings of the Edinburgh Conference, 1966. Edited by J. Lyons and R. J. Wales. Edinburgh University Press. 1966. Pp. 243. 42s.

This book reports the proceedings of a conference held in Edinburgh in March 1966. There were five invited papers; by J. P. Thorne, R. J. Wales and J. C. Marshall, David McNeill, J. Fodor and M. Garrett, E. S. Klima and Ursula Bellugi. Each paper had prepared comments from chief discussants and was open to discussion by the audience of invited guests. In condensing this discussion and in producing the book before the conference became history the editors have been most efficient.

Thorne's paper serves as an introduction to the whole book, several of his topics being taken up by others later. McNeill, and Klima and Bellugi cater more for the specialist in language acquisition, providing many examples of the child's developing syntactic system. The remaining two papers are of the most general interest.

Wales and Marshall provide a synopsis and review of recent psycholinguistic experiments. In the main these experiments explore the way the human brain implements the syntactic and semantic constraints upon utterances, described by the systems of linguists. Thus the authors have to devote some space to describing these systems. That is the current brief in psycholinguistics but Wales and Marshall go beyond it in three important ways: (1) They show that such *passé* fields as word-association and nonsense-syllable learning can be related to contemporary psycholinguistics. (2) They give appropriate importance to the semantic component in any plausible model of linguistic performance, rejecting the notion that to understand a sentence means to assign to it a correct structural description. They might have taken this point further to advocate that the syntactic and semantic were only separable for convenience's sake and suggested experiments on inter-relationships between them. (3) While accepting an innate predisposition in the human for learning certain universal attributes of language structure, and rejecting Markov-type learning models, they urge that such temporary acceptance of language as *sui generis* should not obscure possible important similarities between language structuring and other behavioural or perceptual structuring. If Wales and Marshall can be criticized it is for including too much in their review and thus hiding these points, which are conspicuously absent in other psycholinguistic metatheory. From the number of experiments summarized and distinctions drawn this paper will undoubtedly become used as a text, and benefit the many with thoroughness rather than the few with elegance.

The elegance is provided by Fodor and Garrett's paper on the distinction between competence and performance models. A summary follows; two independent senses are described in which the distinction has been used (*a*) as between descriptive system and utterance—the abstract and the observable and (*b*) as between the abstract system and the interaction of psychological variables (memory capacity, etc.) with the neural embodiment of the information that system describes—both unobservable. Fodor and Garrett concentrate on the second distinction and review a number of experiments designed simply to test the adequacy of the descriptive system without any special modifications, as a model for the analysis of sentences by listeners in perceptual and memory tasks. The negative results obtained seem to rule out a simple analysis-by-synthesis model based on transformational grammar. However for long-term memory, it appears that what is

stored is a form approximating the base structure. (The sentences *John is easy to please* and *John is eager to please* share a surface structure but differ in base structure.). Thus linguistic units in the base structure appear to be psychologically real, but rules mapping them into the surface structure may not be. The relationship between the competence model and the performance model is thus revealed by experiments to be an abstract one. In concluding discussion Garrett reports a recent development by the MIT linguist Haj Ross which suggests that successive experiments will not necessarily show this abstractness to be extreme; Ross modified part of the grammar in a way that both incorporated some discrepant psychological results and elegantly tied up a separate set of inadequacies in the grammar.

Fodor and Garrett's paper gave rise to the most stimulating discussant contribution of the conference. The philosopher L. J. Cohen proposed a new theory of the listener as distinct from the speaker based on the abandonment of any *complete* grammatical analysis as a performance model and giving greater importance to semantic variables. The conceptual complexity of this paper was rather high (although the sentences were short) and unfortunately the audience failed to appreciate it, virtually ignoring it in discussion. This nicely proved Cohen's own point about the importance of the semantic component in the listener model.

In short this book presents some lively and useful psychosyntactics and psychosemantics. What, then, of psychopragmatics, psychophonology and psychophonetics? These potential sub-disciplines had their advocates in the general discussion; they are bound eventually to benefit from the vogue for psychological studies of grammar, and in omitting such fields the conference did not misrepresent current research activity. The fact that the phonetic vowel chart in the jacket design is disarranged is presumably sheer artistic licence rather than a Freudian exclusion of the substance of language from the subject matter of linguistics.

MARK P. HAGGARD.

On Human Communication. By Colin Cherry. 2nd Edition. London and Cambridge, Mass.: M.I.T. Press. 1966. Pp. xiv + 337. 80s.

Professor Cherry's book *On Human Communication*, first published in 1957 (and reviewed in detail by W. E. Hick in this Journal, 1958, **10**, 111) has been sufficiently influential and widely read in the past 10 years to be termed a classic. Since then so much has happened in the communication sciences that it would have entailed writing a new and different book to cover the new information, much of it very technical. Professor Cherry has preferred instead to keep the original work intact and to make only a few minor additions. The book will continue to be a useful, illuminating and lively reference for those entering the field, but psychologists who possess the first edition need not rush to exchange it for the second.

ANNE TREISMAN.

Psychological Principles Appropriate to Social and Clinical Problems. By John C. Raven. London: H. K. Lewis. 1966. Pp. xx + 187. 25s.

It is not often we get a chance to see an eminent psychologist as naked and unashamed as Raven allows us to see him here. In a sense this is why I would not recommend the book to undergraduate students and other innocents as a basic introduction to the field of clinical psychology: such a reader would not be in a position to evaluate the idiosyncrasies. This may be the brand image of clinical psychology for many but for at least this clinical psychologist it is not the generic image of the craft. For a reader who wants to know something of the Crichton school and for young clinical psychologists who may be getting prematurely set in their ways the book has a good deal to offer. The level of the offering is variable: at times elementary precepts are set out, at times hints of promising lines of development are given; appropriate bibliographical detail is sometimes given, sometimes witheld. At times sceptically, at times innocently, the author questions what tests really get at. One gets a good view of the process of "comparative matching," and we are also given a good deal of metaphysics about ego boundaries, intentions and time.

P. R. F. CLARKE.

MEMBERS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY

Honorary Member

Sir Frederic BARTLETT, C.B.E., F.R.S., St. John's College, Cambridge.

Ordinary Members

J. ANNETT, D.Phil., Department of Psychology, The University, Hull.
S. M. ANSTIS, Ph.D., Department of Psychology, The University, Bristol 8.
M. ARGYLE, M.A., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
H. ASHER, B.A., Department of Physiology, University of Birmingham, Birmingham 15.
Professor R. J. AUDLEY, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
B. BABINGTON SMITH, O.B.E., M.A., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
A. D. BADDELEY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
Professor P. BAKAN, Ph.D., Department of Psychology, Michigan State University, East Lansing, Michigan, U.S.A.
Professor H. B. BARLOW, M.D., M.B., B.Chir., University of California, School of Optometry, Berkeley 4, California, U.S.A.
Mrs. E. BELBIN, Ph.D., 8 Lansdowne Road, Cambridge.
A. J. BENSON, M.B., Ch.B., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.
Professor D. E. BERLYNE, Ph.D., Department of Psychology, University of Toronto, Toronto 5, Canada.
Professor P. BERTELSON, Ph.D., Laboratoire de Psychologie, Université Libre de Bruxelles, 115 Avenue Adolphe Buyl, Bruxelles, Belgium.
C. B. BLAKEMORE, B.Sc., Department of Psychology, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5.
G. S. BRINDLEY, M.D., F.R.S., Trinity College, Cambridge.
D. E. BROADBENT, Sc.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
Professor P. L. BROADHURST, Ph.D., Department of Psychology, University of Birmingham, Birmingham 15.
Professor J. BROWN, Ph.D., Department of Psychology, The University, 7 Sydenham Terrace, Newcastle upon Tyne 2.
D. J. BRUCE, Ph.D., Department of Education, 5 Salisbury Villas, Station Road, Cambridge.
P. E. BRYANT, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
A. W. H. BUFFERY, Ph.D., Unit for Research on Medical Applications of Psychology, 5 Salisbury Villas, Station Road, Cambridge.
F. W. CAMPBELL, M.B., Ch.B., Ph.D., D.O.M.S., Physiological Laboratory, Downing Place, Cambridge.
Miss V. R. CANE, M.A., 14 Little St. Mary's Lane, Cambridge.
A. CARPENTER, M.B., B.Chir., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
Professor E. C. CHERRY, D.Sc., Department of Electrical Engineering, Imperial College of Science and Technology, Exhibition Road, London, S.W.7.
M. B. CLOWES, Ph.D., C.S.I.R.O., Computing Research Section, P.O. Box 109, Canberra City, A.C.T., Australia.
W. P. COLQUHOUN, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
K. J. CONNOLLY, B.Sc., Department of Psychology, The University, Sheffield 10.
R. CONRAD, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
D. W. J. CORCORAN, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
A. COWEY, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
J. J. COWLEY, M.Sc., Department of Psychology, The University, Aberdeen.
F. I. M. CRAIK, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

Miss A. CRAWFORD, Ph.D., Department of Psychology, The University, 7 Abercrombie Square, Liverpool 7.

Professor E. R. F. W. CROSSMAN, Ph.D., Department of Industrial Engineering, University of California, Berkeley 4, California, U.S.A.

S. CROWN, Ph.D., 5 Northwick Avenue, Kenton, Middlesex.

H. C. A. DALE, Ph.D., Department of Psychology, The University, Hull.

Professor K. DANZIGER, D.Phil., Department of Psychology, York University, Keele Street, Toronto 12, Canada.

R. DAVIS, D.Phil., Department of Psychology, The University, Reading, Berks.

B. DELISLE BURNS, M.R.C.S., L.R.C.P., Division of Physiology and Pharmacology, National Institute for Medical Research, Mill Hill, London, N.W.7.

Professor J. A. DEUTSCH, D.Phil., Department of Psychology, University of California, San Diego, La Jolla, California, U.S.A.

N. F. DIXON, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

P. C. DODWELL, Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.

J. DREVER, M.A., Principal, University of Dundee, College Gate, Dundee, Scotland.

Professor G. C. DREW, M.A., Department of Psychology, University College, Gower Street, London, W.C.1.

Professor J. T. EAYRS, D.Sc., Department of Neuro-endocrinology, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5.

A. ELITHORN, M.D., M.B., B.Chir., M.R.C.P., D.P.M., Royal Free Hospital, Lawn Road, Hampstead, London, N.W.3.

G. ETTLINGER, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5.

Professor B. M. FOSS, M.A., Institute of Education, Malet Street, London, W.C.1.

Professor E. FRASER, Ph.D., Department of Psychology, The University, Aberdeen.

J. L. GEDYE, M.B., B.Chir., 1 Royston Road, Harston, Cambridge.

F. H. GEORGE, Ph.D., Department of Psychology, The University, Bristol 8.

K. GIBBINS, Ph.D., Department of Psychology, The University, 7 Sydenham Terrace, Newcastle upon Tyne 2.

C. B. GIBBS, Ph.D., Control Systems Laboratory, Division of Mechanical Engineering, National Research Council, Ottawa, Ontario, Canada.

Mrs. F. GOLDMAN-EISLER, D.Phil., Psycholinguistics Research Unit, Department of Phonetics, University College, Gower Street, London, W.C.1.

G. W. GRANGER, Ph.D., Department of Psychology, 7 Kepier Terrace, Gilesgate, Durham.

J. A. GRAY, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.

A. H. GREGORY, Ph.D., Department of Psychology, The University, Manchester 13.

Professor R. L. GREGORY, M.A., Department of Machine Intelligence and Perception, Bionics Research Laboratory, The University, Forrest Road, Edinburgh 1.

Professor R. A. M. GREGSON, Ph.D., Department of Psychology and Sociology, University of Canterbury, P.O. Box 1471, Christchurch, C.I., New Zealand.

Professor S. GRIEW, Ph.D., Department of Psychology, Otago University, Dunedin, New Zealand.

G. C. GRINDLEY, M.A., Psychological Laboratory, Downing Street, Cambridge.

A. M. HALLIDAY, M.B., Ch.B., M.R.C. Neurological Unit, National Hospital, Queen Square, London, W.C.1.

M. S. HALLIDAY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.

M. HAMMERTON, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.

R. HARPER, Ph.D., 18 Leng Crescent, Norwich.

A. D. HARRIS, M.R.C.S., L.R.C.P., Littlemore Hospital, Littlemore, nr. Oxford.

Miss A. W. HEIM, Ph.D., Psychological Laboratory, Downing Street, Cambridge.

L. J. HERBERG, M.B., B.Ch., D.Phil., Institute of Neurology, National Hospital, Queen Square, London, W.C.1.

Mrs. B. M. HERMELIN, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5.

W. E. HICK, M.D., M.B., 66 Bishops Road, Trumpington, Cambridge.

R. HINCHCLIFFE, M.D., Institute of Laryngology and Otology, Royal National Hospital for Throat, Nose and Ear, Grays Inn Road, London, W.C.1.

Professor R. A. HINDE, D.Phil., Sub-Department of Animal Behaviour, High Street, Madingley, Cambridge.

D. H. HOLDING, Ph.D., Department of Psychology, The University, Gandy Street, Exeter.

G. HORN, M.B., Ch.B., School of Anatomy, The University, Cambridge.

N. HOTOPH, M.A., London School of Economics, Houghton Street, Aldwych, London, W.C.2.

I. P. HOWARD, Ph.D., Department of Psychology, York University, Keele Street, Toronto 12, Canada.

Professor C. I. HOWARTH, D.Phil., Department of Psychology, The University, Nottingham.

J. A. M. HOWE, Ph.D., Department of Machine Intelligence and Perception, Bionics Research Laboratory, The University, Forrest Road, Edinburgh 1.

D. H. HOWES, Ph.D., Room E6-121 (Neurology), Boston Veterans Administrative Hospital, 150 South Huntington Avenue, Boston, Massachusetts 02130, U.S.A.

I. M. HUGHES, Ph.D., Department of Psychology, Taylor Building, The University, Aberdeen.

Professor I. M. L. HUNTER, D.Phil., Department of Psychology, The University, Keele, Staffs.

Professor H. M. B. HURWITZ, Ph.D., Department of Psychology, University of Tennessee, Knoxville, Tennessee 37916, U.S.A.

Miss H. IKEDA, Ph.D., Department of Experimental Ophthalmology, Institute of Ophthalmology, Judd Street, London, W.C.1.

J. G. INGHAM, Ph.D., Social Psychiatry Research Unit, Llandough Hospital, Penarth, Glamorgan.

Mrs. S. D. IVERSEN, Ph.D., Psychological Laboratory, Downing Street, Cambridge.

Professor M. A. JEEVES, Ph.D., Department of Psychology, University of Adelaide, G.P.O. Box 198D, Adelaide, S. Australia.

A. R. JONCKHEERE, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

Mrs. D. JOYCE, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

R. B. JOYNSON, M.A., Department of Psychology, The University, Nottingham.

Professor H. KAY, Ph.D., Department of Psychology, The University, Sheffield 10.

R. P. KELVIN, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

Mrs. M. O. KERR, B.A., Royal Free Hospital, North Western Branch, Lawn Road, London, N.W.3.

C. C. KIERNAN, B.A., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

M. KINSBOURNE, B.M., B.Ch., M.R.C.P., Institute of Experimental Psychology, 1 South Parks Road, Oxford.

D. R. J. LAMING, Ph.D., Psychological Laboratory, Downing Street, Cambridge.

D. LEGGE, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

J. A. LEONARD, Ph.D., Department of Psychology, The University, Nottingham.

B. N. LEWIS, B.Sc., 20 Hill Rise, Richmond Surrey.

N. E. LOVELESS, Ph.D., Department of Psychology, St. Andrew's University, Queen's College, Dundee.

Professor D. M. MACKAY, Ph.D., Department of Communications, The University, Keele, Staffs.

N. J. S. M. MACKINTOSH, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.

N. H. MACKWORTH, M.B., Ch.B., Ph.D., 429 Atlantic Street, Stamford, Conn., U.S.A.

Miss F. E. MACNEILL, Ph.D., Department of Psychology, The University, Oxford Road, Manchester 13.

Mrs. I. MARTIN, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5.

Miss S. MAYNE, Ph.D., 88 Bulmershe Road, Reading, Berks.

R. MELZACK, Ph.D., Department of Psychology, McGill University, Montreal, Canada.

Professor T. R. MILES, M.A., Department of Psychology, University College of North Wales, Bangor.

Mrs. B. MILNER, Ph.D., Montreal Neurological Institute, 3801 University Street, Montreal 2, Canada.
N. MORAY, Ph.D., Department of Psychology, The University, Sheffield 10.
J. MORTON, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
G. H. MOWBRAY, Ph.D., Applied Physics Laboratory, Johns Hopkins University, Silver Spring, Maryland, U.S.A.
N. MROSOVSKY, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
W. R. A. MUNTZ, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
D. J. MURRAY, Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.
L. J. NEWSON, Ph.D., Department of Psychology, The University, Nottingham.
N. O'CONNOR, Ph.D., Social Psychiatry Unit, De Crespigny Park, Denmark Hill, London, S.E.5.
R. C. OLDFIELD, M.A., M.R.C. Speech and Communication Research Unit, The University, 31 Buccleuch Place, Edinburgh 8.
Professor I. OSWALD, M.D., M.B., B.Chir., Department of Psychological Medicine, University of Western Australia, Nedlands, Western Australia.
J. M. OXBURY, Ph.D., Neurology Department, Churchill Hospital, Oxford.
Miss M. PICKERSGILL, Ph.D., Department of Psychology, Bedford College, Regent's Park, London, N.W.1.
Professor R. W. PICKFORD, D.Litt., Ph.D., Department of Psychology, University of Glasgow, Glasgow, W.2.
M. F. PIERCY, M.A., Maida Vale Hospital for Nervous Diseases, Edgware Road, London, W.9.
I. PLEYDELL-PEARCE, Ph.D., Department of Psychology, The University, Bristol 8.
E. C. POULTON, M.B., B.Chir., M.R.C.S., L.R.C.P., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
K. A. PROVINS, Ph.D., Department of Psychology, University of Adelaide, Adelaide, Australia.
P. M. A. RABBITT, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
Professor R. L. REID, M.A., Psychological Laboratory, The University, Gandy Street, Exeter.
Miss H. E. ROSS, Ph.D., Department of Psychology, The University, Hull.
W. A. H. RUSHTON, Ph.D., Sc.D., M.R.C.S., L.R.C.P., F.R.S., Trinity College, Cambridge.
I. S. RUSSELL, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
Professor R. W. RUSSELL, Ph.D., D.Sc., Department of Psychology, Indiana University, Bloomington, Indiana, U.S.A.
Professor D. RUSSELL DAVIS, M.D., M.B., B.Chir., M.R.C.P., M.R.C.S., L.R.C.P., D.P.M., Department of Mental Health, Bristol Royal Infirmary, Bristol 2.
B. SEMEONOFF, Ph.D., Department of Psychology, The University, 60 The Pleasance, Edinburgh 8.
T. SHALLICE, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
M. E. SIME, B.A., Department of Psychology, The University, Sheffield 10.
Professor W. SLUCKIN, Ph.D., Department of Psychology, The University, Leicester.
J. R. SMYTHIES, M.Sc., M.D., M.B., B.Chir., M.R.C.P., D.P.M., University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh 10.
A. C. STANILAND, Ph.D., Department of Psychology, The University, Gandy Street, Exeter.
Miss H. STEINBERG, Ph.D., Department of Pharmacology, University College, Gower Street, London, W.C.1.
A. W. STILL, Ph.D., Department of Psychology, The University, 7 Kepier Terrace, Gilesgate, Durham.
H. C. W. STOCKBRIDGE, M.A., Quest, Tilford, Farnham, Surrey.
R. G. A. STRETCH, Ph.D., Department of Psychology, University of Alberta, Edmonton, Alberta, Canada.
Professor A. SUMMERFIELD, B.Sc., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

Professor N. S. SUTHERLAND, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
Professor J. R. SYMONS, Ph.D., Psychology Department, Kings College, Old Aberdeen.
Professor J. SZAFRAN, Ph.D., Department of Psychology, University of Southern California, University Park, Los Angeles, California 90007, U.S.A.
Professor H. TAJFEL, M.A., Department of Psychology, The University, Bristol 8.
D. H. TAYLOR, Ph.D., Psychology Department, The University, Southampton.
Mrs. B. TIZARD, Ph.D., 83 Burbage Road, London, S.E.24.
Mrs. A. TREISMAN, D.Phil., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
Professor M. TREISMAN, M.B., B.Ch., Department of Psychology, The University, Reading, Berks.
Professor E. L. TRIST, O.B.E., M.A., Graduate School of Business Administration (Management Area), University of California, Los Angeles, California 90024, U.S.A.
Professor A. M. UTTLEY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
Professor P. H. VENABLES, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
Professor M. D. VERNON, M.A., Sc.D., Department of Psychology, The University, Reading, Berks.
Miss M. A. VINCE, M.A., Psychological Laboratory, Downing Street, Cambridge.
D. M. VOWLES, D.Phil., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
G. K. WALLACE, Ph.D., Department of Psychology, The University, Reading, Berks.
P. B. WARR, Ph.D., Department of Psychology, The University, Sheffield 10.
Mrs. E. WARRINGTON, Ph.D., Psychological Department, National Hospital, Queen Square, London, W.C.1.
P. C. WASON, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
A. J. WATSON, M.A., Psychological Laboratory, Downing Street, Cambridge.
R. H. J. WATSON, Ph.D., Department of Nutrition, Queen Elizabeth College, Campden Hill Road, London, W.8.
Miss K. P. WATTS, Psychological Laboratory, Downing Street, Cambridge.
Professor L. WEISKRANTZ, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
A. T. WELFORD, Sc.D., Psychological Laboratory, Downing Street, Cambridge.
T. C. D. WHITESIDE, M.B.E., Ph.D., M.B., Ch.B., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.
J. W. WHITFIELD, M.A., Department of Psychology, University College, Gower Street, London, W.C.1.
P. WHITTLE, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
R. T. WILKINSON, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
The Hon. Mrs. M. WILLIAMS, Ph.D., Addenbrooke's Hospital, Lensfield Road, Cambridge.
J. A. WILSON, B.Sc., 32 Cotham Park, Bristol 6.
Mrs. M. WYKE, Ph.D., Maida Vale Hospital for Nervous Diseases, Edgware Road, London, W.9.
Professor O. L. ZANGWILL, M.A., Psychological Laboratory, Downing Street, Cambridge.

REVUE CANADIENNE CANADIAN JOURNAL DE PSYCHOLOGIE OF PSYCHOLOGY

VOLUME 41, No. 2 and 3 APRIL/JULY 1967

Published for the Canadian Psychological Association by
PUBLIÉ POUR LA SOCIÉTÉ CANADIENNE DE PSYCHOLOGIE PAR
UNIVERSITY OF TORONTO PRESS

THE CANADIAN JOURNAL OF PSYCHOLOGY is published bi-monthly. Annual subscriptions, $10.00, single number $2.00.

Contributions. Original manuscripts and correspondence on editorial matters should be sent to:

The Editor, Canadian Journal of Psychology, McMaster University, Hamilton, Ontario.

Subscriptions to Journal. Orders and correspondence regarding subscriptions, changes of address, and purchase of back numbers should be sent to:

Canadian Psychological Association,
The Secretary-Treasurer,
Suite 210, 225 Lisgar Street,
Ottawa 4, Ontario.

THE QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

Vol. XIX 1967 Part IV

THE SUPPRESSION OF VISUALIZATION BY READING

BY

L. R. BROOKS

Department of Psychology, McMaster University, Ontario

Four experiments are described which demonstrate a conflict between reading verbal messages and imagining the spatial relations described by those messages. Listening to the same messages did not produce comparable interference with visualization. The conflict between reading and visualizing was obtained even when the subject previously had seen the referent of the message. In contrast, when the subject was induced to treat the messages as rote strings of words instead of visualizing their referents, reading was a more effective means of presentation than was listening.

Two interpretations of these results were proposed. (*a*) Visualization and reading compete for the use of neural pathways specialized for visual perception. (*b*) The process of reading hinders the conversion of input material into any non-verbal form; that is, reading forces the subject to deal with information in a more exclusively verbal form than does listening. It was suggested that regardless of interpretation this method provides a means of investigating the internal processes underlying concrete verbal reference.

Introduction

Assume that one is given a verbal description of the layout of the streets in a town or the positions of furniture in another room. An immediate response to these descriptions would probably be to imagine how the layout of those items might look. If called upon to act on this information or to repeat the description to someone else, it is likely that one would refer to his visualization of the information while responding. Introspectively, at least, the notion that this type of internal representation involves the visual or visual-motor systems is a compelling one. Even if one does not experience visual images, the phrases "imagine what it would look like," and "visualize the relations" seem to describe some part of this mental process.

The hypothesis that this type of internal representation involves the visual system has been persistent in psychology. The early associationist and introspective psychologists assigned an important role in thinking to visual images, which were considered to be at least a partial revival of the immediate effects of visual perception. Since the ascendance of behavioural methods, numerous psychologists have searched for objective indices which implicated the visual and visual-motor systems in visual imagery. Eye movements, for example, have been correlated with thinking about visual tasks (Deckert, 1964; Lorens and Darrow, 1962). Studies of this type have been collected and set in the context of the peripheralist–centralist controversy by McGuigan (1966). Arguments supporting the importance of orientation movements, such as eye movements, for covert representations of information have been presented by Berlyne (1965).

The research in this paper uses gross performance measures rather than EEG or eye-movement recordings as a means of implicating visual activity in thinking.

Before describing this approach, the model implicit in the opening paragraph will be made more explicit. A subject is presented with a verbal message which describes a set of spatial relations. Concurrent with receiving the message, he generates an internal representation which involves some portion of the visual or visual-motor systems. If the subject is required to repeat or act on the message, the internal representation is again generated and serves as one determiner of output. The visual or visual-motor systems are thus partially occupied during both input of the message and output tasks which use the information in the message.

If the verbal message itself were written, then the subject would be required to use his visual system simultaneously for two separate purposes: namely, to represent covertly the spatial relations and to identify the written words. If there is some overlap in the mechanisms which are used to accomplish these two tasks then the subject ought to experience a conflict which he would not experience if he *listened* to the message. Similarly, if the act of output involves looking at words, for example inspecting a written passage for errors or writing down the verbal message, then a conflict ought to be produced which would not be present if the output involved only listening or speaking. Experiments 1 and 2 are designed to test these two implications.

Experiment 1

Subjects were presented with a series of messages which described spatial relations. Some of these messages were spoken to the subject; others were spoken at the same rate but were accompanied by the simultaneous exposure of a typewritten copy of the message. The subjects were asked after each message to repeat it verbatim. According to the above model, reading the message should lead to less accurate repetition than listening to the message, since listening alone would not conflict with the use of the visual system for visualization of the spatial relations. However, for this evidence to be pertinent, it also must be shown that there is not some peculiarity of the method of presentation which biases against reading *per se*. To demonstrate this, nonsense messages were presented which had the same sentence form and length and which would therefore have the same presentation characteristics as the spatial messages. If there were no difference between performance after reading and performance after listening with this nonsense material, then the characteristics of the physical presentation could be eliminated as a sufficient explanation for the results with the spatial material.

Method

Material. Figure 1 shows an example of a spatial message, the matrix which it describes and a nonsense message derived from it. The first sentence of each spatial message was always the same, and the same square in the four-by-four matrix was always designated as the starting square. The numbers were always the digits one to eight and were always mentioned in consecutive numerical order. The only way in which spatial messages differed was the sequence of transitions (up, down, right, left) from one square to another. These sequences were designed so that two different digits were never assigned to the same square by a message, and that a digit was never placed outside the matrix. Messages used for practice were reduced to only the first four sentences.

The nonsense material was derived from the spatial material by replacing the words *right*, *left*, *up*, *down*, with the words *quick*, *slow*, *good*, *bad*. In this way, material was obtained that is identical to the spatial material in length and form of presentation, and which involves strings of polar opposites with the same sequential restraints. However, for the present experiment (but not for Experiment 2), it was necessary to introduce another difference between the two types of material. Pilot subjects performed so poorly with the eight sentence nonsense messages that differentiating the presentation conditions would have been impossible. Consequently, only five sentences were used in each nonsense

FIGURE 1

		3	4
	1	2	5
		7	6
		8	

A sample of the experimental material.

Spatial material	Nonsense material
In the starting square put a 1.	In the starting square put a 1.
In the next square to the *right* put a 2.	In the next square to the *quick* put a 2.
In the next square *up* put a 3.	In the next square to the *good* put a 3.
In the next square to the *right* put a 4.	In the next square to the *quick* put a 4.
In the next square *down* put a 5.	In the next square to the *bad* put a 5.
In the next square *down* put a 6.	In the next square to the *bad* put a 6.
In the next square to the *left* put a 7.	In the next square to the *slow* put a 7.
In the next square *down* put an 8.	In the next square to the *bad* put an 8.

message. The spatial messages could not also be reduced to five sentences without both eliminating the need to visualize and resulting in so few errors that once again the presentation conditions could not be differentiated.

Design. The two types of material, spatial and nonsense, were each presented in two ways: (1) *Listening* (*L*) the message (set of sentences) was played from a tape recorder. (2) *Listening and reading* (*LR*) a series of 3 × 5 index cards, each with one sentence of the message typed on it, were presented concurrently with the tape. The listening and reading condition was chosen over a reading alone condition so that differences in performance could be attributed to the presence of absence of the visual presentation and not to the presence or absence of the auditory presentation. In addition, the LR condition minimized the effect of individual differences in reading speed. Subjects reported a strong tendency to read each word in time with the auditory presentation, which tended to insure that even the fast readers did not have time to glance away from the card before the presentation of the next sentence.

For each of the experimental conditions, the task was simply to repeat the message verbatim. Each condition was given two times to each subject. The spatial and the nonsense trials were run in separate blocks; the material in each block was presented in one of the following sequences: L, LR, LR, L or LR, L, L, LR. Both the order of blocks of material and the order of types of presentation within these blocks were counterbalanced across subjects.

In summary, the experimental design was two-by-two within-subjects, one variable being spatial or nonsense material, and the other variable being listening (L) or listening plus reading (LR) presentation.

Subjects. The subjects were eight summer school students who were serving to fulfil a course requirement in introductory psychology. Subjects were tested individually.

Procedure. For the initial practice period, each subject was shown a diagram similar to that in Figure 1, and listened to the message that went with it. He was told that he would be asked to either read or listen to a set of similar messages which would *not* be accompanied by a diagram, and that immediately after each message was presented he was to repeat it word for word. He was then given a practice series of four, four-sentence spatial messages without visual matrices; two of these messages presented by listening and two by listening and reading. This practice block was sufficient to encompass most of the very large practice effect which occurs with this material At the conclusion of the practice block, the subject was given either the block of spatial trials or the block of nonsense

trials. The block of nonsense trials was preceded by the information that four meaningless adjectives were going to be substituted into the sentences but that the task was still to repeat the message verbatim. He was given two, four-sentence nonsense messages for practice and then the block of nonsense trials.

The subjects were asked to be sure to keep their eyes on the cards during the reading trials and were reminded of this on the infrequent occasions that they glanced away. The messages in all conditions were presented at the rate of approximately one sentence per 1·5 sec. No specific knowledge of results was given, although all subjects were periodically encouraged about their performance.

Results

The subjects uniformly reported that they performed the spatial task by "picturing the pattern formed by the numbers," and then "reading from the pattern" when repeating the message. The nonsense task was reported to have been performed by trying to retain the sequence of adjectives alone, and then, in output, reinserting them into the grammatical context.

TABLE I

AVERAGE ERRORS PER MESSAGE

	Spatial material	*Nonsense material*
Listening	1·2	2·3
Listening and reading	2·8	1·3

The average errors out of seven possible for the spatial conditions, and four possible for the nonsense conditions are shown in Table I. A within-subjects analysis of variance showed the material by presentation interaction significant at the 0·05 level ($F(1, 7) = 6{\cdot}67$) but neither of the main effects significant (material $F(1, 7) < 1{\cdot}0$; presentation $F(1, 7) = 4{\cdot}73$). Seven of the eight subjects made more errors after reading the spatial messages than listening to them. Seven of the eight subjects made more errors after listening to the nonsense material than after reading it.

It can be concluded that the listening–reading condition (LR) hindered the repetition of the spatial material but did not hinder the repetition of the nonsense material. This result is consistent with the hypothesis that reading interferes with the generation of an internal representation of the spatial relations. The performance with the nonsense material indicates that this effect is not due to some feature of presentation which made reading inherently more difficult.

EXPERIMENT 2

The purpose of this experiment was to demonstrate the conflict between reading and visualizing during output instead of input as in Experiment 1. Subjects learned a message to a criterion of one verbatim repetition. They then were asked either to say the key words of the message (the spatial relations themselves) or to underline those same key words on a multiple choice test (a written copy of the message which listed the four possible options at each choice point). It was predicted that saying the spatial relations would be more rapid and involve fewer long pauses than would underlining. To underline the written spatial relations the subject would have to read them, and reading should conflict with his visualization of the spatial relations.

The same procedure also was run with the nonsense materials described in Experiment 1. This was designed to control for the possibility that the sheer movements involved in underlining take more time than does speaking, regardless of visualization.

Method

Design. The same two types of material were used, with the modification that each of the nonsense as well as each of the spatial messages consisted of eight sentences. Each message was taught to a criterion of one verbatim repetition in order to obtain differences in output that were not due to different levels of acquisition of the two types of material.

After criterion was reached on a given message, the subject was requested to give one of two types of output (1) *spoken*: the key words alone were spoken by the subject (right, up, right, down, down, etc.; or, quick, good, quick, bad, bad, etc.). (2) *underlining*: a sheet of paper was uncovered in front of the subject on which each of the sentences in the messages was typed, one under the other with a double typewriter space between. All of the four possible key words were bracketted within each sentence. The subject's task was to proceed down the page, underlining the appropriate word in each sentence—much in the manner of taking a multiple-choice test.

Each of the four conditions were given twice to all subjects. One trial was run on each of the conditions before the second trial in any condition was administered. The sequence in which the four conditions were run was counterbalanced across subjects. In summary, the experimental design was two-by-two within-subjects, with one factor being spatial or nonsense material and the other factor being underlining or spoken output.

Subjects. The subjects were eight first-year university students, who were participating to fulfil a course requirement in introductory psychology.

Procedure. Each subject was introduced to the material in the same way as in the previous experiment. He then was given two practice trials, one with a spatial message and one with a nonsense message, and was asked to give both types of output for each message. The instructions for the mode of output required on a particular trial was given by saying either "spoken" or "written" within 5 sec. after informing the subject that he had repeated the message correctly in acquisition. The initial response times and the inter-response times were recorded on a Gerbrands event recorder, driven at 1 cm./sec., by depressing a telegraph key each time the subject either said or marked a key word.

To facilitate original acquisition all messages were presented at a slower rate than in the previous experiment: approximately one sentence per 2·5 sec. rather than one per 1·5 sec. As in the first experiment, no specific knowledge of results was given regarding output performance.

Results

Seven of the eight subjects reported that they "pictured the pattern" described by the spatial material, and that they referred to this pattern during output. All subjects reported that they learned the nonsense material by noticing sequential patterns in the key words, and that at output these patterns of words "just came."

The average number of trials to acquisition of each message was 1·2 for the spatial material, and 2·0 for the nonsense material. Seven of the eight subjects made more errors on the nonsense material than on the spatial material, and the eighth subject made the same number on each type of material. The subjects claimed that being able to figure out a spatial pattern for the spatial material helped considerably in acquisition.

TABLE II

AVERAGE PERFORMANCE TIME IN SECONDS

	Spatial material	*Nonsense material*
Speaking	7·6	8·0
Underlining	15·5	11·1

The average total output time per message is shown in Table II. A two-by-two within-subjects analysis of variance on these times showed the main effect due to type of output significant with a probability of less than 0·01 ($F(1, 7) = 14·51$) and

the interaction of output and material significant with a probability of less than 0·05 ($F(1, 7) = 5·74$). The main effect due to material was not significant ($F(1, 7) = 2·59$). Thus, the relative difficulty of the written output differed for the two types of material.

An informal observation was made that the subjects glanced away from the page occasionally during underlining of the spatial material. When this was pointed out after the experiment, the subjects explained the glancing away as "getting the pattern back." It seemed that this phenomenon could be demonstrated by plotting the average inter-response times sequentially through a trial and looking for a peak which would indicate that the subject had glanced away. However, the exact inter-response time during which a subject glanced away was so variable that the average curve was essentially flat for inter-response times three through six. Another approach to measuring the phenomenon is to look for skewness in the distribution of inter-response times on each trial. If there are in fact one or two abnormally long inter-response times, then a skewness statistic should show the effect. The statistic chosen for this purpose was the mean minus the median. The average deviation of the mean from the median was 0·71 for written spatial, 0·26 for spoken spatial, 0·18 for written nonsense, and 0·21 for spoken nonsense. Comparisons made by t-test for correlated means showed written spatial to be different at 0·05 from all other groups, with no differences among the other groups.

The combination of these two findings, i.e. the interaction in total output times and the skewness of the individual inter-response time distributions, supports the conclusion that the subjects found conflict between internally representing the matrix and underlining the key words, and that they solved the conflict by taking relatively long pauses during the output. This conflict is not inherent to the underlining mode of output as is shown by the lack of similar effect for the output of the nonsense material.

Experiments 1 and 2 together support the conclusion that, for this situation, reading suppresses spatial organization. Both the performance data and the reports of the subjects indicate a conflict between reading and imagining the matrices described by the spatial messages. Two interpretations of this result can be made. The first interpretation, *specific visual involvement*, is that which was mentioned in the introduction: namely, that both reading and internal representation of spatial organizations utilize some of the mechanisms specialized for handling visual perception. The second interpretation, *general overload*, of the conflict between reading and spatial organization rests on the assumption that both reading and spatial organization are more complicated than the tasks with which they were contrasted in these experiments: namely, listening and sequential organization. The subject attempts to treat reading and visualizing successively in order to reduce an overload on his *general* processing capacity. In contrast, the notion of specific visual involvement attributes the conflict to competition for the same *visual* processing mechanisms. Experiments 3 and 4 were designed to obtain information more critical to the specific visual involvement interpretation.

Experiment 3

Two considerations influenced the design of this experiment. (*a*) The plausibility of the general overload explanation would be weakened if it could be shown that a conflict still existed after the subject had already seen the matrix. It would seem reasonable that if the matrix had been seen, then imagining it while reading would be a good deal less taxing and conflict due to overload should disappear. (*b*) The assumption has been made that the conflict between reading and the spatial task is occasioned by the process of visualization, rather than by the material itself (or, more

precisely, by the immediate semantic response to the spatial words). Substance could be added to this argument if the conflict were eliminated when the subjects treated the spatial messages as a sequence of words rather than as a description to be visualized.

The design of Experiment 3 is as follows. The initial, spoken presentation of each message is accompanied by one of two displays; either a matrix showing the numbers in the appropriate squares, or a written listing of the key words alone (right, up, right, down, etc.). After each of these initial displays, the verbal message alone is presented either by listening, or by listening and reading, as in Experiment 1. When the subject treats the message as describing a matrix, then reading should interfere with the visualization and thus slightly delay the subsequent repetition of the message. However, when the subject treats the message as a sequence of words, then reading should have a slightly facilitating effect, as it did with the nonsense material in Experiment 1.

Method

Design. The spatial material from Experiment 1 was used, and all details of the verbal presentation and output were the same as for the spatial material in Experiment 1. However, prior to the L or LR presentation of the message, each subject was given an L presentation of the same message together with either (*a*) a 5 in. × 8 in. card on which was drawn a four-by-four matrix containing the numbers in the locations described by the message, or (*b*) a 5 in. × 8 in. card on which each of the seven key words (directions describing the transitions from one number to another) were typed in vertical sequence. Thus, on each trial there was a "set presentation" (matrix or word sequence), followed by a "verbal presentation" (listening or listening and reading, as defined in the first experiment), followed by a verbatim repetition of the verbal message by the subject.

Each of the four conditions were given two times to each subject. One trial was run with each of the conditions before the second trial in any condition was administered. The sequence in which the four conditions were run was counterbalanced across subjects. In summary, the experimental design was two-by-two within-subjects, with one variable being matrix or word sequence as the set presentation and the other variable being L or LR verbal presentation.

Subjects. The subjects were eight first-year university students, who were participating to fulfil a course requirement in introductory psychology.

Procedure. Each student was introduced to the material in the same way as in the previous experiment. He was then shown the two types of initial presentation and was told that it was of the utmost importance that he deal with the materials in the manner that they were given. That is, for the word sequence condition he was not to try to picture the referent of the words, but rather to deal with them as if they were simply an arbitrary verbal sequence. He was then given two practice trials on each of the types of material. To emphasize the instructions he was asked on the practice trials using the matrix to fill in a blank matrix after he had given a verbatim repetition, and on the practice trials using the word sequence he was asked to give the string of key words alone as rapidly as possible. On all subsequent trials, he was asked to give only the verbatim repetition.

Results

The average time to give a verbatim repetition of the message is given in Table III for each of the experimental conditions. The anticipated result was found; output after reading (LR) took longer than output after listening (L), but only when the subject was asked to visualize the matrix described by the message. A two-by-two factorial within-subjects analysis of variance on the average times for each subject for each condition showed the main effect due to set presentation significant with a probability less than 0·01 ($F(1, 7) = 15{\cdot}72$) and the interaction of verbal presentation and set presentation significant with a probability less than 0·05 ($F(1, 7) = 6{\cdot}12$). The main effect due to verbal presentation was not significant ($F(1, 7) = 1{\cdot}23$).

The simple effect of L *vs.* LR was tested for both levels of the set presentation and found to be significant at 0·05 for both levels.

TABLE III

AVERAGE PERFORMANCE TIME IN SECONDS

	Matrix	*Word sequence*
Listening	12·5	22·5
Listening and reading	14·0	17·5

The average number of errors per repetition was not as useful in this experiment as it was in Experiment 1, since five of the eight subjects made no errors after a matrix presentation. However, the error data for the sequence presentation are consistent with the time data given above; six subjects in this condition showed more errors after L than LR verbal presentation, one subject showed no errors in either, and the final subject reversed the trend by making one error after LR and none after L. The ceiling effect evident in these data is presumably due to having two presentations of each message (one set presentation and one verbal presentation) prior to the verbatim repetition, as well as the additional visual display during the set presentation.

This result for the time data and the tendency in the same direction for the error data demonstrate that the reading conflict can be varied by setting the subjects to treat the same verbal material in two different ways. In addition, it demonstrates that the conflict is not limited to a situation in which the subject has to deduce the referent of the verbal message.

EXPERIMENT 4

This experiment was designed to provide the same type of information as Experiment 3 and, in addition, to resolve an ambiguity specific to the interpretation of Experiment 2. In Experiment 2 an attempt was made to show that when a person visualized during output, his performance was hindered on a mode of output which entailed reading. To demonstrate this, two types of output were compared: speaking the key words and underlining the correct words in a written passage. However, there is an important difference between these two output modes; speaking does not require attending to anything outside of the responder, while underlining requires reading the words on the page. It is possible that the results were due to the necessity of attending to external stimuli rather than to the fact that reading specifically involves the visual system. The present, error detection, experiment was designed to eliminate this possibility.

Subjects are taught a message to a criterion of one verbatim repetition, as in Experiment 2. The subjects' task is to detect a possible error in a subsequent presentation of a sequence of three key words taken from the message. The major variable in the experiment is the manner in which the error messages are presented. After some spatial messages, the three key words are spoken rapidly to the subject; after other messages, the key words are displayed on an index card. It is expected that error detection will be slower with the written presentation. However, as in Experiment 3, this result is expected only when the subject is visualizing the material. Consequently, the error detection task is performed both when the subject is instructed to visualize the matrix and when he is given the spatial relations as a rote verbal sequence.

Method

Design. There were three phases to each trial: acquisition, error message presentation, and judgement. One acquisition condition was identical to the matrix condition of the last experiment; the subjects were shown the relevant matrix while they listened to a spatial message. In the other acquisition condition, the "word sequence" condition, the subject listened to a rapid presentation of the key words of a spatial message. Since it was important for interpretation of the error detection data that the subjects have the message fluently, they were given presentations until they could repeat the material without hesitation.

Each error message consisted of three words and was either spoken or typewritten. To generate an error message, a sequence of three successive key words (spatial relations) were taken from anywhere in the message except the initial three. On 50 per cent. of the messages an error was introduced by changing one of the words to its polar opposite (right to left, up to down, and vice versa). The resulting error is "plausible" in the sense that it would not describe a matrix in which two numbers were assigned to the same square (as would happen, for example, with the sequence "right, left"). A seemingly more straightforward way of running an error detection experiment would be to simply repeat the whole message with or without an error in it. However, when the experiment is performed in this manner the subjects simply generate the sequence of key words in slight anticipation of each word in the error presentation and are thus able to detect the error immediately and virtually without mistakes. By taking three key words from anywhere in the message, the subject is forced to receive the three words and then search the internal representation. This process forces a measurable delay prior to the subject giving his judgement.

In summary, the experimental design was two-by-two within-subjects, with one variable being matrix or word sequence initial presentation and the other variable being written or spoken error message presentation.

Subjects. The subjects were eight first-year university students, who were participating to fulfil a course requirement in introductory psychology.

Procedure. Each subject was introduced to the material in the same way as in the previous experiment. He was then shown a series of error messages while the demonstration matrix was still present. A total of six trials were given in the matrix presentation condition. After each of these six messages were learned a total of four error messages were presented in random order for judgement; correct written, correct spoken, incorrect written, incorrect spoken. The first two of these trials were treated as practice trials and the final four were used for data. Subjects were then given an identical six trials with the word sequence presentation condition.

As in the previous experiment subjects were cautioned to treat the material in the manner given; that is, not to visualize the word sequence material. Whenever a subject made an incorrect judgement on an error message, he was told that he was incorrect and was asked to check the message again. Before each error message was presented, the subject was told to run over the matrix or sequence to himself, and then to give a signal that he was ready. If the subject indicated that he was uncertain about the matrix or sequence he was given another acquisition trial.

Results

The average time from the beginning of a presentation of an error message until the subject's response is shown for each experimental condition in Table IV. As was anticipated, when the material was visualized, time to give judgement was longer with a written error message than with spoken. The same is not true when the material was retained as a verbal sequence. A two-by-two within-subjects analysis

TABLE IV

AVERAGE JUDGEMENT TIME IN SECONDS

	Matrix	*Word sequence*
Listening	4·4	3·4
Reading	5·0	3·1

of variance on the average judgement time per subject per condition showed a significant effect due to material ($F(1, 7) = 17 \cdot 44$) and material by presentation interaction ($F(1, 7) = 7 \cdot 13$) but not presentation ($F < 1 \cdot 0$). Both of the simple effects due to presentation were significant. Six of the eight subjects showed faster judgement times after listening than after reading when the material was visualized; seven subjects showed the reverse effect for rote verbal retention.

This experiment shows that reading during the course of output hinders visualization of material that had previously been learned. Unlike Experiment 2, the contrasting output condition entails listening rather than merely speaking from memory. Thus, the effect cannot be due to merely the necessity of taking in *any* information in the course of performance. This experiment also shows, as did Experiment 3, that the reading–visualizing conflict can be obtained even when the material being visualized had been seen previously.

General Discussion

These experiments demonstrate that in this situation reading competes with the internal representation of spatial information. Both the performance data and the reports of the subjects indicate a conflict between reading and imagining the matrices described by the verbal messages. The effect can be obtained when the messages describe either matrices with which the subject is unfamiliar or matrices the subject has learned or seen previously. The effect can be abolished by inducing the subject to treat the messages as rote sequences of words.

At the present time the most persuasive interpretation of this reading–visualization conflict is that the internal representation of spatial material use mechanisms specialized for *visual* perception. However, it is clear that this interpretation is tentative. It could easily be true that the important thing about the visualization carried out in the present experiment is not its sensory modality but rather that it is organized quite differently from the verbal message itself. That is, it is possible that reading interferes with re-organization of the information from the form given in perception, regardless of what the referent is. Evidence which would select between this interpretation and the specific visual involvement hypothesis could be obtained in a situation in which the same verbal message described characteristics of two different referents, one of which is spatially organized and the other of which is not. If the relative efficacy of reading and listening changed with the referent then one would have a strong argument that the basis of the reading conflict was more specific than simply re-organization from perception.

Regardless of interpretation, the subjects in these experiments did report that they "pictured" the patterns described by the messages. However, when they were presented with a list of features that they would have seen if they had actually inspected a matrix (e.g. outline of the four-by-four matrix, lines between the squares, the digit being described at the moment, other digits than the one being described at the moment), they ascribed to "seeing" anything in only the most gingerly fashion. All of the subjects except one asserted with some conviction that the non-verbal portion of their representation of the information was important to their performance, but only two subjects reported anything which would correspond to a clear image. These two subjects were indistinguishable from the others in speed and accuracy of performance, as was expected from many previous findings on both sensory and thought imagery (Fernald, 1912; review in Oswald, 1962). This result, however, should not be interpreted as indicating that internal representation of visual properties is either entirely verbal or unrelated to the visual system. The 23 subjects who reported picturing the matrix maintained that this activity was not just verbal and

did not back down from this position (although some were in an evident quandary) after they denied actually "seeing" many of the details which would be obvious in vision.

If one accepted the subjects' reports that the way in which they dealt with the matrices was "visual," this still would not be sufficient to establish that visualization used specifically the visual system. The phenomenal property "visual" could be more a result of the subjects' knowledge of what is being described than a result of the part of the nervous system involved in the internal representation. The basic device in this paper for establishing the type of processing is still the conflict with other visual stimulation.

This suggests a final point. If the specific visual involvement interpretation is in fact applicable, it should be possible to show that attending to *any* visual stimulus hinders concurrent visualization. There are two reasons why this research used written descriptive passages instead of other types of visual stimuli to provide interference for visualization. (1) If the stimulus being used for interference were irrelevant to the information being visualized, then it would be likely that an attentional limitation would be met which would force the subject to attend to either the stimulus or the visualization but not both (that is, he would be unable to "think of two things at once"). This would not be bothersome except that the same limitation would probably be met regardless of the modality in which the interfering stimulus were delivered. This situation thus would be inappropriate for demonstrating that visualizing was a specifically *visual* function. (2) One of the most interesting possibilities suggested by the reading conflict is the investigation of the relations between verbal and non-verbal components of internal representation. If reading suppresses some types of visually-mediated organizations, then one has a technique for momentarily changing the relative availability of the verbal and the visualized components of the same thinking process. Using this technique, one could evaluate the subject's relative dependence on verbal and non-verbal components of thought in various operations on different types of material. This type of evidence could form an interesting attack on the psychological problem of concrete verbal reference.

References

Berlyne, D. E. (1965). *Structure and Direction in Thinking*. New York: Wiley.

Deckert, G. (1964). Pursuit eye movements in the absence of a moving stimulus. *Science*, **143**, 1192–3.

Fernald, M. R. (1912). The diagnosis of mental imagery. *Psychol. Monogr.*, **14**, No. 58.

Lorens, S. A., and Darrow, C. W. (1962). Eye movements, EEG, GSR, and EKG during mental multiplication. *Electroenceph. clin. Neurophysiol.*, **14**, 739–46.

McGuigan, F. J. (1966). *Thinking: Studies of Covert Language Processes*. New York: Appleton Century Crofts.

Oswald, I. (1962). *Sleeping and Waking*. New York: Elsevier.

Manuscript received 18th May, 1967.

SOME RELEVANT FACTORS IN THE TRANSFER OF MATERIAL FROM SHORT-TERM TO LONG-TERM MEMORY

BY

RONALD L. COHEN and BO S. JOHANSSON

From the Psychological Laboratory, University of Uppsala, Sweden

The relationship between short-term memory (STM) and long-term memory (LTM) for digits was investigated by means of a Hebb-type experiment, viz. a presentation of a series of nine-digit numbers, in which a certain number recurs at intervals. Improvement in performance, with regard to the critical, or recurring, number was found when the rehearsal grouping was three-three-and-three, but was absent when there were no opportunities for rehearsal or when the rehearsal strategy was grouping five-and-four or searching for systematic numbers.

However, if an overt recall response was given on each occasion the recurring number was presented, improved performance was found even with rehearsal strategy five-and-four.

The conclusion was drawn that rehearsal is the main transferring mechanism from STM to LTM, with the occurrence of an overt recall response as a subsidiary factor.

Introduction

The point of departure for the experiments in this report is the two-mechanism theory of memory by Hebb (1949), and Hebb's (1961) and Melton's (1963*a*) experimental attempts at testing the theory by testing whether performance in repeating a nine-digit number improved when this number recurred at intervals in a series of other such numbers. They did not find any support for the concept of an activity trace which leaves no permanent trace in the CNS, and Melton concluded that perhaps all percepts leave structural traces in the CNS.

Cohen and Johansson (1967) carried out a modified Hebb-type experiment in which they varied the type of response to the recurring or critical number. They found that an overt recall response to the critical number was necessary in order to obtain improvement in performance. That is, improvement was effected during the last phase of the memorizing process (overt recall response) rather than during either of the two earlier phases, viz. stimulation or presentation of the stimulus sequence by the tape recorder, or silent repetition of the stimulus sequence during or after presentation. In the absence of an overt recall response the memorizing process produced activity traces which either left no structural traces in the CNS or else relatively weak structural traces incapable of giving improvement in performance under the conditions used.

The primary aim of the present report was to examine in more detail the role played by the rehearsal phase of the memorizing process in producing improvement in performance. The rehearsal variable has always been difficult to control, and, in fact one of the main criticisms against decay theory is the failure of decay theorists to control or manipulate rehearsal (Postman, 1964). As a consequence, the view of the role of rehearsal differs among the workers in the field. Brown (1958) takes the view that rehearsal has no strengthening effects on the memory trace, but only prevents its decay. This assumption has experimental support from e.g. Brown (1958) and Conrad (1960). Broadbent (1958) would appear to agree with Brown.

Among those reporting the opposite result are Sanders (1961), who has found strengthening effects of rehearsal. This non-agreement may possibly be due to the

al than the other workers. Sanders's
cidental learning, where frequency of
ound to be an important variable in
1962, 1964).
ement in the question of the role of
m mentioned when STM is defined.
isize the same factors in their opera-
e stimulus sequence and recall within
Mackworth, 1964; Melton, 1963a).
s noted that some subjects reported
mental session was to find an effective
d that rehearsal in threes is optimal,
ulus sequence is nine. It was there-
tematically in order to find out how
t of rehearsal in producing improve-
ne-digit number.
orted here was to study the effect of
n improvement in performance.

I

wed only during presentation, in the
nd after presentation and in the third
entation with an overt recall response
earsal strategies were used: grouping
ect of both strategies was measured
ond strategy was investigated under

and Johansson, 1967). A series of 21
corder. The rate of presentation was
umber. After each presentation there
t gave his response, if required. In the
ouncements of the form "that was the
een the numbers when no response was
ter-number interval was empty.
terval was empty regardless of whether

er (the critical number) recurred at the
series. The total series of 21 numbers
17 numbers, and a test series comprising
rred five times in the pre-test series and
rval between the pre-test and the test

ing experimental condition only in that
osition 20. The 3rd, 6th, 9th, 12th and
r random numbers, so that no number
art from this the series of numbers used

training college served as subjects, thus
e been large, most subjects being in the

The rehearsal instructions

Rehearsal in threes. The subjects were instructed to listen carefully to the first three digits, repeat all three silently, and then concentrate on the second three digits. After they were played back, they were to repeat them silently and at the same time do a quick check of the first three if possible. Then they were to concentrate on the last three after which they were to repeat the digits silently or aloud depending on the instructions, or to cease thinking about the number if no response was required.

Rehearsal five-and-four. The subjects were instructed to listen carefully to the first five digits, repeat all five silently, and then concentrate on the last four. When the whole number had been presented, their next action depended on the instructions, as above.

The rehearsal instructions were supplemented with the information that the rehearsal method was of importance for the outcome of the experiment, and that it was of the utmost importance that the subjects should keep to the required method.

All instructions ended with the experimenter encouraging the subjects to concentrate always on the whole nine-digit number and try to report nine digits each time a recall response was to be given, thus avoiding the confounding effects found under two of the conditions in the authors' former report, where the subjects concentrated on remembering only one or two of the digits. When a written response was required, the subjects were instructed to write down the digits in the same order as they had been presented.

Condition A

The subjects were tested individually under this condition and their recall responses were given verbally.

This condition is a repeat of a condition in the authors' former investigation, where the subjects were instructed to listen to, and try to retain each number played back, but to give an overt recall response only when ordered to do so, this being when the experimenter raised his arm. The subjects had to try to retain each number since the signal for recall was not given until the end of the presentation of each number. Recall responses were required only to non-critical numbers (numbers in the 1st, 4th, 8th, 10th, 13th and 16th positions) in the pre-test series and to all four numbers in the test series. There were inter-number announcements after the numbers not requiring a recall response. The 40-sec. pause was used to inform the subjects that they were to reproduce all the numbers which were to follow.

The responses made to the critical number in the pre-test series, were assumed to be only in the form of rehearsal during presentation. Attempts at rehearsal after presentation were discouraged, firstly by the instructions, in which the subjects were told that if the signal for a recall response was not given they could simply forget the number and rest during the 11-sec. interval. And secondly, there were inter-number announcements after all numbers not requiring a recall response which could be expected to interfere with any post-presentation attempts at rehearsal.

This condition was run with two groups of subjects, one being instructed to rehearse in threes (A_{333}) during presentation and the other five-and-four (A_{54}).

Condition B

This condition was run as a group experiment with the subjects writing down their recall responses.

The subjects were instructed to listen to each number played back and to try to memorize it. In the pre-test series, immediately after the nine digits in a number had been played back, the subjects were to repeat the number silently in order to count the number of digits they could repeat in order, and then write down the number, i.e. how many digits they thought they could repeat in order, but not the actual digits. Before the test series began, they were instructed to write down the digits of each number as soon as it had been presented.

During the pre-test series the subjects were assumed to rehearse the critical number during presentation, and to repeat the number silently once after presentation. All subjects did report an estimate of how many digits they could remember in order for all numbers in the pre-test series, and this is taken as indicating that the requirements of condition were met.

This condition was run with two groups of subjects, one being instructed to rehearse in threes during presentation (B_{333}) and the other five-and-four (B_{54}).

Condition C

The subjects were tested individually under this condition and their recall responses were given verbally.

This condition is a repetition of one of the conditions in the authors' previous investigation. Under that condition the subjects were to repeat aloud all the numbers in the pre-test series and the test series. Since improvement in performance was found under this condition it was taken for granted that rehearsal in threes would give a significant improvement in performance. Therefore only the instruction to rehearse five-and-four was investigated. The 40-sec. pause between pre-test and test series was filled with the experimenter chatting with the subject.

Under this condition (C_{54}) the subjects were assumed to rehearse each number during presentation and to give an overt recall response immediately after presentation.

Scoring

In the previous investigation, responses were scored by counting the number of digits correctly reported in order, up to the point of the first error made (after Hebb, 1961). In the present study, the counting was done both forwards from the beginning of the response and backwards from the end of it, this method having the advantage of reducing and normalizing the variance. This revision of the scoring method does not give rise to any changes in the results obtained in the previous investigation.

One method of measuring the effect of the experimental treatment is that of comparing the mean score of the critical number in the 20th position in an experimental group with the score on the corresponding number in a control group. In this case the effect may be obscured by large individual differences in memory span.

One method of overcoming this is to make the comparison within each subject in an experimental group, in order to see whether performance on the critical number in position 20 differed from that on surrounding numbers, viz. those in positions 19 and 21. If his performance is better on the critical number then it can be concluded that this is due either to the recurrence of the number five times before in the series, or else to the critical number being easier than the surrounding numbers.

Five control groups (total N = 96) were run, corresponding to experimental groups A_{54}, B_{54}, B_{333} (of Exp. I), D and E (of Exp. II). The differences between the mean score on the critical number and that on the surrounding numbers ranged from $+0.59$ to -0.54. In no case did this difference come anywhere near significance (the highest obtained t-value was 1·04), so it can be concluded that the critical number did not differ in ease of memorability from the surrounding numbers in this experimental situation. Therefore any differences between the critical number and its surrounding numbers with an experimental group can be attributed to the occurrence of the critical number five times previously in the series. (In view of the unequivocal results obtained with the five control conditions above, it was considered unnecessary to run the controls corresponding to experimental conditions A_{333} and C_{54}, since the same series of numbers was used under all conditions).

It should be noted that the method of giving the recall response (verbally or in writing) is not a variable in this type of experiment. This can be seen by comparing the mean score for numbers 19, 20 and 21 for control condition A_{54} (verbal response) with that for those numbers for control B_{54} (written response), 5·07 and 4·86 respectively.

Results and discussion

Whether significant improvement in performance will take place or not seems to depend on both strategy of rehearsal and amount of rehearsal in the pre-test series. In conditions A and B, where there was rehearsal during presentation only, or rehearsal during presentation together with one silent rehearsal after presentation, respectively, a significant improvement in performance is found only where the rehearsal strategy was grouping in threes. However, when the subjects produced an overt recall response to every stimulus sequence, an improved performance was found even with grouping five-and-four.

The results are interpreted as supporting the theory that rehearsal transfers a stimulus item from a short-term store to a more permanent one. As regards the connection between rehearsal strategy and improved performance, the following explanation is proposed. Firstly, it is assumed that if improvement in performance is to take

TABLE I

GROUP MEAN DIFFERENCE BETWEEN THE CRITICAL NUMBER AND THE AVERAGE OF THE SURROUNDING NUMBERS [20 − (19 + 21)/2], AND THE MEAN NUMBER OF DIGITS CORRECTLY REPORTED FOR THE CRITICAL NUMBER (20)

Condition	$Mean_{20}$	*Diff.* $20 - \frac{19+21}{2}$	$SD_{diff.}$	n	t	p
A_{333}	5·55	1·10	2·06	20	2·39	0·05
A_{54}	4·00	−0·09	2·24	16	—	n.s.
B_{333}	6·42	1·67	1·74	21	4·29	0·001
B_{54}	5·41	0·36	2·22	22	0·74	n.s.
C_{54}	4·83	1·08	2·02	18	2·25	0·05

place, there must be a certain degree of correspondence between the original stimulus sequence and the rehearsal trace. Secondly, Wickelgren (1965) has shown that with too large rehearsal groups, the items in a series tend to change positions. Thus, when rehearsing five-and-four, the rehearsed stimulus sequence may become too different from the original stimulus sequence for an improvement in performance to take place when the stimulus sequence recurs the next time. But, when rehearsing in threes the item positions are better preserved and improvement in performance has a chance to take place the next time the stimulus sequence recurs. As can be seen from Table I, the memory span for both rehearsal strategies is about the same, thus indicating that the strategy of rehearsal is more important than number of digits correctly reported, for improvement in performance to take place.

In the presence of overt recall responses to the critical number (condition C) improvement in performance takes place even when rehearsing five-and-four. The results of the above experiment therefore show the importance of not only a good rehearsal strategy, but also of an overt recall response. This latter point is in line with results reported by Murray (1965), who found that STM for visually presented sequences was better when these were read aloud by the subject during presentation, than when they were read silently.

EXPERIMENT II

The previous experiment showed that either rehearsal during presentation (three-and-three) or rehearsal both during and after presentation resulted in improvement in performance. When testing the decay theory it is important to have both kinds of rehearsal under control, and earlier studies are not free from criticism on this point. For example, Brown (1958) and Peterson and Peterson (1959) controlled rehearsal only in the interval between presentation and recall. When varying the rate of presentation, there is usually no control of rehearsal during presentation (e.g. Mackworth, 1962; Pollack, Johnson and Knaff, 1959) and sometimes not even in the interval between presentation and recall (Waugh and Norman, 1965), unless the recall is paced and is to start immediately after presentation (Conrad, 1957; Conrad and Hille, 1958). Murdock (1965) has proposed the use of a subsidiary task to prevent rehearsal during presentation, and Broadbent (1957) has used a similar technique to prevent rehearsal both during and after presentation. Perhaps the most effective way of preventing rehearsal during presentation, however, is to increase the rate of presentation of the stimulus sequence (Posner, 1963). Therefore, in condition D an attempt was made to prevent rehearsal during presentation by presenting the digits

at a rate of 4 digits/sec., and a recall response was demanded immediately after each presentation of a sequence, thereby preventing rehearsal after presentation.

Recoding is generally assumed to facilitate learning and retention (Miller, 1956; Postman, 1963) and some workers have tried to apply recoding principles to performance in STM (e.g. Melton 1963*a*, *b*). The aim of condition E was to examine whether it was possible to learn the recurring number by a recoding type of rehearsal.

Condition D

The subjects were tested individually under this condition and their recall responses were made verbally.

In the previous experiment the rate of presentation was 1 digit/sec., thus giving the subjects an opportunity of rehearsing the items during presentation. In the present condition, the rate was 4 digits/sec. during the pre-test series, but kept at 1 digit/sec. during the test series. The original length of the inter-number interval was maintained. In order to prevent rehearsal after presentation the subjects were instructed to recall each number as soon as it had been played back. The 40-sec. interval before the change in rate of presentation was filled with the experimenter informing the subject of the change in rate of presentation. No grouping instructions were given.

Under this condition the subjects were assumed to recall the digits of the critical number in the pre-test series aloud, with no rehearsal intervening.

Condition E

This condition was run as a group experiment with the subjects writing down their recall responses.

The instructions under this condition were to listen carefully to what was played back on the tape recorder and try to find out whether each number was arranged according to a system or not. A system was defined as a non-random order of the digits. The numbers in the 1st, 4th, 5th, 8th, 11th, 13th and 16th positions had their digits systematically ordered. The position in the series of each number was announced shortly before its presentation, and the subjects had the 11-sec. interval between numbers in which to write down whether the previous number was systematic or not. After the pre-test series, the subjects were told to write down each number immediately after its presentation. All numbers were read at a rate of 1 digit/sec.

Now the subjects had not only to rehearse the digits but also to try to compare them with information already stored in the LTM in order to find out whether they were arranged systematically. (E.g. the sequence 321 is systematic for a subject who has already learned the sequence 321 or 123, to take two possibilities. It can be tested for systematization only by comparing it with previous learned sequences or rules stored in LTM.) This should then give the critical number a good chance of leaving a permanent trace in the LTM. The activity of the subjects during the pre-test series could be controlled since they had to indicate the numbers in which they found systems. The subjects detected a mean of 6·9 out of 7 numbers which were systematic and reported a mean of only 1·6 wrong detections. These data are taken to indicate that the subjects followed instructions.

Results and discussion

The result of condition D indicates that where rehearsal during and after presentation is prevented no improvement in performance occurs. A possible criticism which could be made here is that whereas the pre-test series used 4 digits/sec., the test series used 1 digit/sec., and that a trace left by sequences presented at 4 digits/sec. may not be detectable when tested with a 1 digit/sec. sequence. The number of digits correctly reported of the number in the 15th position was 5·46, and this does not differ significantly from the number of digits correctly reported of the surrounding numbers (14 and 16). (A control showed that the corresponding difference under condition C is significant at the 0·05 level with $t = 2{\cdot}17$, $d.f. = 17$). Therefore, it seems clear from the above that no improvement occurred under condition D regardless of whether this is tested using a presentation rate of 1 digit/sec. (the 20th number)

or one of 4 digits/sec. (the 15th number). Thus, the conclusion from the previous experiment that rehearsal has the key role in transferring stimulus-items from STM to LTM gains further support.

TABLE II

GROUP MEAN DIFFERENCE BETWEEN THE CRITICAL NUMBER AND THE AVERAGE OF THE SURROUNDING NUMBERS [20 − (19 + 21)/2], AND THE MEAN NUMBER OF DIGITS CORRECTLY REPORTED FOR THE CRITICAL NUMBER (20)

Condition	*Mean*$_{20}$	*Diff.* $20 - \frac{19+21}{2}$	$SD_{diff.}$	n	t	p
D	5·26	−0·27	3·26	16	—	n.s.
E	5·60	0·53	2·30	20	1·00	n.s.

In condition E no significant improvement in performance was found. The subjects were asked after completion of the experiment, if they had noticed a recurrence of a number or not. Eleven out of 20 in the experimental group and 12 out of 19 in the control group answered *yes* to the question. However, on asking the subjects to report the number they thought had recurred, only three out of the 11 subjects, who had answered *yes* in the experimental group, were able to report enough of the critical number to make it identifiable. Since there were approximately as many in the experimental group as in the control group who reported a recurrence, and since the majority of the experimental subjects made an erroneous identification, this result is taken as indicating that no detectable trace of the critical number had been left in the CNS under this condition.

The reason for the lack of improvement can be attributed to the rehearsal strategy. The subjects were asked to rehearse each stimulus sequence in order to find out whether there was a system in the sequence or not, and this probably resulted in a trace being left which bore only a slight resemblance to the original number.

This interpretation would mean that if there are any facilitating effects of an encoding mode of rehearsal in recall, then there must be some system or structure in the sequence being encoded.

Summary and conclusions

The experiments comprising the present report have shown that a dichotomy, improvement in performance/no improvement in performance, can be obtained using the present experimental set-up, depending on the conditions used. This does not necessarily mean that one condition gives rise to a structural trace whereas another condition gives rise to an activity trace only, but rather that those conditions favouring an improvement in performance give rise to a structural trace which is strong enough to withstand interference and/or decay, whereas those conditions not favouring an improvement give rise to, at most, a structural trace which is comparatively weak.

The results can then be summarized as follows:

(1) No structural trace could be detected if no rehearsal took place (condition D).

(2) Allowing rehearsal during presentation, or both during and after presentation, results in a trace having structural properties (conditions A and B).

(3) Introducing an overt recall response seems to result in a stronger structural trace than if only rehearsal is allowed (condition C).

(4) A general finding of this report is that there must be a certain degree of correspondence between the structural trace produced by rehearsal and the original stimulus sequence if improvement in performance is to take place.

There may, of course, be other ways of producing structural traces than by rehearsal, but the efforts in this report to produce improvement in performance without rehearsal have not resulted in any success. This seems to support the view that rehearsal is the most important mechanism for producing structural changes in the CNS.

In the previous study the rate of presentation was 1 digit/sec. and although rehearsal was not controlled by giving instructions, many subjects reported that they had tried to group at least some of the digits in each sequence. In this case the giving of an overt recall response would be critical since such a rehearsal strategy is anything but optimal, and the conditions used in the latter study are probably very comparable with conditions B_{54} and C in the present one.

Thus with an optimal rehearsal strategy (three-three-and-three in the present case) an overt response is unnecessary, but with a stretegy less than optimal (five-and-four, for example) such a response is necessary for improved performance. Further, an overt recall response, without rehearsal grouping, is not sufficient to give improved performance.

These results also mean that a description of the measures taken to control rehearsal should be included together with the specifications of the number of presentations of the stimulus and the time interval between presentation and recall, when an experiment is operationally defined as a STM or a LTM experiment. An experimental design with one presentation of the stimulus sequence, and recall within a short time of presentation, may be in danger of bringing in LTM if ample opportunities for rehearsal are given.

This view of STM differs from that of Brown and Broadbent as regards the effect of rehearsal in so far as they assumed that rehearsal merely postponed the onset of the decay. Broadbent (1963) does distinguish between a pre- and a post-rehearsal STM-system on the basis of the findings reported by Sperling (1960) on STM in vision, but he assumes that both STM-systems follow the same principles.

When defending a one-mechanism theory of memory, Melton (1963*b*) argues that he cannot understand why "an encoded event with a frequency of 1 should, when lost, have its loss ascribed to decay, while an encoded event with a frequency of 2, 3, . . ., *n* should have its loss attributed to interference." The answer to Melton's argument from the point of view presented in the present report should be that if "encoded" means the same as "rehearsed," then the forgetting of the event should obey the laws of interference, whereas if the event is not rehearsed, then its forgetting should most probably obey the laws of decay, although the possibility of a more permanent storage even without rehearsal cannot be excluded (cf. Waugh and Norman, 1965, who assume that events become permanently stored even without rehearsal).

This investigation was supported by the Swedish Council for Social Science Research.

References

Broadbent, D. E. (1957). Immediate memory and simultaneous stimuli. *Quart. J. exp. Psychol.*, **9**, 1–11.

Broadbent, D. E. (1958). *Perception and Communication.* New York: Pergamon Press.

Broadbent, D. E. (1963). Flow of information within the organism. *J. verb. Learn. verb. Behav.*, **2**, 34–9.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. exp. Psychol.*, **10**, 12–21.

COHEN, R. L., and JOHANSSON, B. S. (1967). The activity trace in immediate memory; a re-evaluation. *J. verb. Learn. verb. Behav.*, **6**, 139–43.
CONRAD, R. (1957). Decay theory and immediate memory. *Nature*, **179**, 831–2.
CONRAD, R. (1960). Very brief delay of immediate recall. *Quart. J. exp. Psychol.*, **12**, 45–7.
CONRAD, R., and HILLE, B. A. (1958). The decay theory of immediate memory. *Canad. J. Psychol.*, **12**, 1–6.
HEBB, D. O. (1949). *The Organization of Behavior*. New York: Wiley.
HEBB, D. O. (1961). Distinctive features of learning in the higher animal. In DELAFRESNAYE, J. F. (Ed.), *Brain Mechanisms and Learning*. Oxford: Blackwell.
MACKWORTH, J. F. (1962). Presentation rate and immediate memory. *Canad. J. Psychol.*, **16**, 42–7.
MACKWORTH, J. F. (1964). Interference and decay in very short-term memory. *J. verb. Learn. verb. Behav.*, **3**, 300–8.
MECHANIC, A. (1962). Effects of orienting task, practice, and incentive on simultaneous incidental and intentional learning. *J. exp. Psychol.*, **64**, 393–9.
MECHANIC, A. (1964). The responses involved in the rote learning of verbal materials. *J. verb. Learn. verb. Behav.*, **3**, 30–6.
MELTON, A. W. (1963*a*). Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.*, **2**, 1–21.
MELTON, A. W. (1963*b*). Comments on Professor Peterson's paper. In COFER, C. N. and MUSGRAVE, B. S. (Eds.). *Verbal Behavior and Learning: Problems and Processes*. New York: McGraw-Hill.
MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, **63**, 81–97.
MURDOCK, B. B. jr. (1965). Effects of a subsidiary task on short-term memory. *Brit. J. Psychol.*, **56**, 413–9.
MURRAY, D. J. (1965). Vocalization-at-presentation and immediate recall, with varying presentation-rates. *Quart. J. exp. Psychol.*, **17**, 47–56.
PETERSON, L. R., and PETERSON, M. J. (1959). Short-term retention of individual verbal items. *J. exp. Psychol.*, **58**, 193–8.
POLLACK, I., JOHNSON, L. B., and KNAFF, P. R. (1959). Running memory span. *J. exp. Psychol.*, **57**, 137–46.
POSNER, M. I. (1963). Immediate memory in sequential tasks. *Psychol. Bull.*, **60**, 333–49.
POSTMAN, L. (1963). Does interference theory predict too much forgetting? *J. verb. Learn. verb. Behav.*, **2**, 40–8.
POSTMAN, L. (1964). Short-term memory and incidental learning. In MELTON, A. W. (Ed.). *Categories of Human Learning*. New York: Academic Press.
SANDERS, A. F. (1961). Rehearsal and recall in immediate memory. *Ergonomics.*, **4**, 25–34.
SPERLING, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.*, **74**, Whole No. 498.
WAUGH, N. C., and NORMAN, D. A. (1965). Primary memory. *Psychol. Rev.*, **72**, 89–104.
WICKELGREN, W. A. (1965). Size of rehearsal group and short-term memory. *J. exp. Psychol.*, **68**, 413–9.

Manuscript received 10th March, 1967.

REDUNDANCY EFFECTS IN SHORT-TERM MEMORY

BY

D. W. J. CORCORAN* and D. L. WEENING

From the U.S. Navy Electronics Laboratory, and San Diego State College, San Diego, California

Five experiments are reported whose purpose was to demonstrate that short-term memory is improved by redundancy within the material. In Experiment I "tunes" containing two, three, four and five tones of differing frequencies had to be coded into digits 1–5, to indicate the order of the pitches in a tune. Performance on stimuli containing correlated amplitude and duration were compared with the uni-dimensional condition. Experiment II repeated I, but required intensity to be coded. Experiment III required pitch coding under three conditions including that when amplitude and frequency were uncorrelated, and compared the performance of musically trained subjects with non-musicians. Experiment IV repeated III, but subjects were informed of the relation between dimensions. Experiment V involved "shadowing" the tunes by whistling simultaneously with the stimulus.

It was concluded (*a*) that intercorrelation improves, but zero correlation impairs short-term memory; (*b*) that knowledge of the relation between dimensions improves performance in the correlated condition, but does not prevent impairment under zero correlation; and (*c*) the performance of musically trained subjects exceeds that of controls and is unaffected by the presence of a correlated or uncorrelated dimension.

Introduction

Eriksen (1952, 1953) and Green and Anderson (1956) have shown that the time required in searching for a given stimulus is shorter if that stimulus differs in two or more dimensions, rather than in a single dimension, from the rest of the population. Eriksen and Hake (1955) and Garner and Creelman (1964) in similar investigations demonstrated that the absolute identification of stimuli varying only in hue, size or brightness was improved if each hue had its characteristic size, each size its characteristic brightness or each brightness its characteristic hue. Further improvements could be shown when all three qualities were intercorrelated. The effects of interdimensional correlation upon recognition in these studies have been considered to be due to the greater informational redundancy in the bi- and tri-dimensional material. Thus interdimensional correlation has been identified as a form of redundancy, acting like repetition or duplication of the same information to improve discrimination (Garner, 1962).

That interdimensional correlation might improve resolution within the memory trace is suggested by the work of Conrad (1964) and especially by Baddeley (1966), who has shown that lists of acoustically similar verbal items are more difficult to recall correctly after a short interval than comparable lists of acoustically dissimilar items. Since acoustically similar items differ from each other in fewer phonetic dimensions than dissimilar items, the improved memory for the dissimilar material may be considered to be due to redundancy within the material. The spoken B (bee), for example, may be considered to be composed of two phonemes, /b/ and /i/. This sound differs from V (vee) by a single phoneme, /b/, and discrimination between the two rests entirely upon resolution of /b/ from /v/. When presented with B and Z, however, the initial phoneme may be forgotten without loss since Z is composed of

* Exchange scientist from M.R.C. Applied Psychology Research Unit, Cambridge, England, where part of the experimental work was conducted.

two redundant phonemes, /ɛ/ and /d/. (This description oversimplifies the phonetic structure of spoken letters, and is used merely as an example.)

The major purpose of the following experiments was to demonstrate more formally the influence of interdimensional correlation on immediate memory and therefore link two previously separate areas of research. In the process of eliminating other possible explanations of the phenomenon another related question was raised: that of the influence of an uncorrelated dimension upon immediate memory.

Experiment I

Method

Stimulus material and procedure. A set of four tape recordings, F, FA, FD and FAD, was made, each containing 11 stimuli. A given stimulus consisted of from two to five separate tones of different frequencies. Tape F began with a 2-tone stimulus made up of a 1 sec. 1,250 cps. tone followed immediately by a 1,000 cps. tone of the same duration. The subject was requested to code this stimulus into numbers, by giving a value of 2 to the higher tone and a value of 1 to the lower. Thus his written answer of 2, 1 simply indicated the order in which the two tones had been presented. After the 2-tone stimulus came a 3-tone stimulus in which a higher frequency tone of 1,500 cps. had been added to the original two. The first 3-tone stimulus consisted of 1 sec. bursts of 1,250, 1,000 and 1,500 cps. tones in that order and the response 213 was checked before the subject was allowed to continue with the task. Nine further stimuli made up the whole of Tape F. These consisted of three more 3-tone stimuli, three 4-tone stimuli and three 5-tone stimuli. The 3-tone stimuli contained the same three tones used in the examples, the 4-tone stimuli contained an additional tone of 1,750 cps. which had to be coded as a 4, and the 5-tone stimuli an additional 2,000 cps. tone, to be coded as a 5. The particular orders were determined by a random procedure with the restriction that within a given stimulus no tone could occur more than once. The nine-test stimuli were as follows: 321, 132, 213, 4231, 3214, 2143, 35142, 45132, 12435.

In Tape FA each frequency had its own characteristic amplitude. Thus the 1,000 cps. tone was always presented at the lowest intensity, the 1,250 cps. tone at the second lowest intensity and so on. The actual pressure levels of the tone at the subject's ears are difficult to specify, since the stimuli were delivered over a loudspeaker. The tones were recorded, however, at values which were approximately equal in interval, as determined by the method of magnitude estimation (Stevens, 1956). The levels at the source (a Peters SPD5 audiometer) were 43 dB., 57 dB., 62 dB., 66 dB. and 69 dB., and the gain on the recording instrument was kept constant. Apart from the introduction of intensity the stimuli of FA were identical to those of Tape F.

In Tape FD each frequency had a different duration. The 1,000 cps. tone was exposed for ½ sec., the 1,250 cps. tone for ¾ sec., the 1,500 cps. tone for 1 sec., the 1,750 cps. tone for 1¼ sec., and the 2,000 cps. tone for 1½ sec. The level at the source was kept constant at 62 dB.

TABLE I

Ordering of Conditions for Each Subject. (Experiment I)

Subject	Order 1	2	3	4
1	F	FAD	FA	FD
2	F	FAD	FD	FA
3	FAD	F	FA	FD
4	FAD	F	FD	FA
5	FA	FD	F	FAD
6	FA	FD	FAD	F
7	FD	FA	F	FAD
8	FD	FA	FAD	F

Both the amplitude and duration of each frequency differed in Tape FAD but otherwise this tape was identical to F, FA and FD. Thus the stimuli in all tapes were composed of the same frequencies in the same order and although FA, FD and FAD contained additional correlated dimensions, all tapes required identical coding.

Subjects were not informed of the relevancy of the redundant dimensions, but were requested before each recording to code in terms of pitch. They were, however, warned that the recording would sound a little different from the previous one. Several subjects noticed variations in intensity; only one commented upon differences in duration and none reported having explicitly used the correlated dimension(s). Subjects were instructed to write nothing until the complete stimuli had been presented.

Subjects. Eight laboratory personnel, six males and two females, served as subjects. All were scientific members of the laboratory, or technical assistants; they were assigned to conditions in the order shown in Table I.

Results and discussion

Since subjects were aware of the digits they would use in their answers, the only uncertainty involved was the *order* in which the digits were to be placed. Accordingly, the measure proposed by Crossman (1960) was used to assess the amount of information transmitted for each stimulus. Crossman proposed that the amount of presented information should be estimated by $\log_2 n!$ (where n is the number of items (tones) within a stimulus) and that the information lost should equal $\log_2 t!$, where t is the number of transposed items. The transmitted information for each stimulus was therefore estimated by the proportion $(\log_2 n! - \log_2 t!)/\log_2 n!$.

TABLE II

MEAN PROPORTION OF TRANSMITTED INFORMATION (PITCH). (EXPERIMENT I)

	Number of tones in stimulus			
Condition	3	4	5	*Overall mean*
F	0·968	0·982	0·828	0·926
FA	1·000	0·991	0·905	0·965
FD	0·968	0·967	0·824	0·920
FAD	0·984	1·000	0·844	0·943

Table II presents the mean proportion of transmitted information per stimulus for each type of stimulus. This table shows (*a*) that only 5-tone stimuli caused any real difficulty and (*b*) that the proportion of information transmitted by Tapes F and FD was less than that transmitted by FA and FAD. The overall scores obtained on F did not differ significantly from those of Tape FD, but were reliably lower than the scores obtained on FA ($p < 0·02$) and FAD ($p < 0·05$), as assessed by the Wilcoxon test. Similarly, the scores on FD were less than those on FA and FAD ($p < 0·02$ and $p < 0·05$, respectively).

Since the coding was kept constant throughout the four conditions, the differences in performance are probably attributable to the ease with which the stimuli were retained during the interval when the coding was taking place. The results therefore indicate that the introduction of intensity as a redundant dimension improved short-term storage, but that the further introduction of duration did not have an additional beneficial effect.

The superiority of FA and FAD over F could, however, have been due to one of four factors: (*a*) the correlated dimension of intensity may have increased the redundancy of the information carried by the tones, thus increasing their discriminability in memory; (*b*) the intensity differences may have been more easily retained than differences in pitch and subjects might have explicitly used the former in *preference* to the latter in the FA and FAD conditions; (*c*) the variations in the intensity may have simply improved the perception of pitch (it is known, for example, that pitch perception is influenced by intensity (Doughty and Garner, 1948)), or (*d*) the redundancy may have improved immediate recognition, not storage.

Experiment II was designed to eliminate the second possibility (*b*) that intensity was used in preference to pitch in Experiment I.

Experiment II

Method

Stimulus material and procedure. Tapes FA and FAD were retained and two additional tapes, A and AD, were constructed. In both new tapes the stimuli were identical to all previous tapes in the way their stimuli had to be coded. The tones on Tape A were all of 1,500 cps. and of 1 sec. duration, but differed in intensity; the intensities used were identical with those of Tapes FA and FAD. On Tape AD all tones were 1,500 cps., but varied in amplitude plus duration; the values on these dimensions were identical with Tapes FA and FAD.

Subjects were requested to code each stimulus into numbers, giving high numbers to high intensities and low numbers to less intense tones. Thus the previous experiment was essentially repeated except that intensity rather than pitch judgements were required.

Subjects and procedure. A further eight laboratory subjects were used in a design identical to that used in Experiment I.

Results and discussion

Table III shows the mean proportion of information transmitted per stimulus for 3-, 4- and 5-tone stimuli under A, FA, AD, and FAD conditions. The results are essentially the same as those of the previous experiment. Although there is some small indication that intensity was more easily coded than pitch, the introduction of frequency differences (on tapes FA and FAD) resulted in superior performance. Overall performance on A and AD did not differ significantly from each other, but FA and FAD yielded reliably greater information transmission than A and AD ($p < 0{\cdot}05$ in both cases).

TABLE III

Mean Proportion of Transmitted Information (Intensity). (Experiment II)

Condition	Number of tones in stimulus			
	3	4	5	*Overall mean*
A	0·984	0·973	0·858	0·938
FA	1·000	1·000	0·873	0·958
AD	1·000	0·917	0·838	0·918
FAD	1·000	0·991	0·877	0·956

The results of Experiment II serve to reject the hypothesis that intensity was coded in preference to pitch in Experiment I. If this had occurred, then (*a*) FA would not have yielded better performance than A, and (*b*) FA in Experiment I would have resulted in about equal performance to A in Experiment II.

The next experiment was designed to eliminate the third hypothesis, (*c*) that the differences in amplitude simply enhanced the perception of pitch.

Experiment III

If the superior performance resulting from the correlation of amplitude and frequency were due simply to the better perception of pitch with varying amplitudes (or the better perception of intensity with varying frequency), then it is of no importance that the two dimensions were correlated. Variations in amplitude which are unrelated to frequency should therefore have the same effect as variations in amplitude *correlated* with frequency. In the following experiment, therefore, pitch was coded when (*a*) only frequency differences existed in the material, (2) when amplitude and frequency were correlated, and (3) when they were uncorrelated. Experiment III also considered another variable, which will be discussed at a later stage and which necessitated a group of subjects who were musically trained.

Method

Stimulus material. Three tapes were constructed, the details of which are presented in Table IV. Tape F contained a 2-tone stimulus and three 3-tone stimuli as practice items; these were followed by four 4-tone, four 5-tone and four 6-tone stimuli. All tones were of 1 sec. duration and recorded at an arbitrary meter reading of −14 dB. The six tones had the frequencies shown in Table IV, which can be seen to be lower than those used in previous experiments. (The lower frequencies were used because it was considered possible that the relatively high frequencies used in the early experiments may have been particularly difficult to rehearse since they were outside normal voice frequencies. This factor did not appear to influence the results and will not be further considered.) The correct coding of the stimuli were: 21, 213, 321, 132, 3142, 4213, 1432, 2314, 15324, 52314, 35241, 41352, 463125, 243651, 652143, 546231.

TABLE IV

Characteristics of Stimuli. (Experiments III and IV)

	All tapes	*FAc*	*FAuc*		
			Amplitude re zero dB.		
Code	*Frequency in cps.*	*Amplitude re zero dB.*	4-*tone*	5-*tone*	6-*tone*
1	700	−35	−28	−14	−7
2	800	−28	−35	−35	−35
3	900	−21	−14	−21	−14
4	1,000	−14	−21	−7	−21
5	1,100	−7	—	−28	0
6	1,200	0	—	—	−28

Tape FAc was identical to F except that each frequency was recorded at the amplitude shown in Table IV. It will be noted that amplitude and frequency are perfectly correlated in FAc, whereas the amplitudes at which the tones were recorded on Tape FAuc (also shown in Table IV) bore no relationship to the frequencies; in fact, the rank order correlations are all zero. The four training items, which were inserted before the test items on all tapes did not vary in amplitude.

Subjects and procedure. Twelve undergraduates of San Diego State College served as subjects. Six were studying music as their principal subject and six were taking other varied courses. Subjects were tested in Music/non-Music pairs, each pair receiving a different order of the six possible orders of three conditions. Stimuli were relayed over earphones such that the amplitudes for the least and most intense tones varied between approximately 60 and 95 dB. re 0·0002 microbar at the ear. Subjects were requested to code in terms of pitch and were not informed of the existence of differing amplitudes. Differences in intensity were usually perceived, to judge from the reports, but usually not as correlated to the pitch.

Results and discussion

Table V presents the mean proportion of transmitted information under each condition for both groups. Within the musical group no differences in performance were evident between conditions. In the non-musical group two of the three comparisons were found to be statistically reliable, but it was necessary to use a parametric technique to achieve satisfactory levels of confidence. Overall performance under FAuc was found to be lower than that under F ($t = 2·58$, $p < 0·05$) and FAc ($t = 2·88$, $p < 0·05$) in the non-musical group. The scores achieved by musicians were reliably higher than for non-musicians under all conditions ($p < 0·01$ in all comparisons).

TABLE V

MEAN PROPORTION OF TRANSMITTED INFORMATION. (EXPERIMENT III)

	No. of tones				
	4	5	6	*Overall mean*	
F	0·995 0·857	0·908 0·734	0·774 0·676	0·892 0·755	Musical non-Musical
FAc	0·989 0·927	0·894 0·794	0·803 0·645	0·895 0·789	Musical non-Musical
FAuc	0·953 0·659	0·915 0·647	0·744 0·465	0·870 0·590	Musical non-Musical

The results of Experiment III were therefore somewhat unexpected (*a*) in failing to confirm the earlier results at a satisfactory level of confidence and (*b*) by showing a significant deterioration in performance in the uncorrelated condition. Whilst the results certainly reject the hypothesis that differences in amplitude simply improved perception of pitch, they also present a new problem: why did performance deteriorate rather than remain unaffected by the uncorrelated dimension? Perhaps the most immediate question generated by Experiment III, however, is whether the results of the two previous experiments were simply chance observations. Experiment IV was designed to eliminate the latter possibility and at the same time throw some light upon the mechanism which is apparently causing deterioration in performance in the uncorrelated condition.

EXPERIMENT IV

Method

Stimulus material, subjects and procedure. The stimuli were identical to those used in Experiment III. The procedure differed in one respect in that subjects were made fully aware of the relationships between amplitude and frequency. It was suggested to them

that when the two parameters were uncorrelated they should attempt to ignore the amplitude differences. A further six State College students were tested, none of whom were musicians. The design was identical to that used in the previous experiment.

Results and discussion

The results, which are shown in Table VI, confirmed the positive findings of all previous experiments. All inter-condition comparisons were reliable, the Wilcoxon test yielding Ts of zero, $p < 0{\cdot}05$ in all cases. Thus knowledge of the relationship between the two dimensions enhanced the beneficial effects of correlated dimensions, but comparisons between the present results and those of Experiment III show that the deterioration in the uncorrelated condition was probably unaffected by the knowledge that amplitude and frequency were uncorrelated. The improvement of 11 per cent from F to FAc in Experiment IV was reliably greater than the 3 per cent. improvement over F shown in Experiment III. The Mann–Whitney test gave $U = 5$, and a two-tailed $p = 0{\cdot}042$. The deterioration under FAuc conditions in Experiment III did not differ reliably from that in Experiment IV, however.

TABLE VI

Mean Proportion of Transmitted Information. (Experiment IV)

	No. of items			
	4	5	6	*Overall mean*
F	0·740	0·575	0·549	0·621
FAc	0·795	0·763	0·641	0·733
FAuc	0·601	0·477	0·456	0·511

The results of Experiments III and IV were analysed further to discover any effects of order of presentation of conditions. Performance under each condition was compared when each had been immediately succeeded by one of the other conditions and when presented as the first condition. The results are summarized in Table VII, which presents the mean transmitted information under each contingency; each mean is based upon six subjects. None of the comparisons are statistically reliable even by a single-tailed test.

TABLE VII

Mean Proportion of Transmitted Information for Each Condition when Preceded by F, FAc, and FAuc. (Experiments III and IV)

	Preceding condition			
Condition	*1st condition*	*After F*	*After FAc*	*After FAuc*
F	0·767	—	0·683	0·804
FAc	0·801	0·856	—	0·758
FAuc	0·552	0·611	0·767	—

Experiment V

The final experiment seeks to test the hypothesis (*d*) that redundancy within the material improved the immediate recognition of the pitch, rather than its storage. The appropriate experiment is of a somewhat different kind from the previous experiments.

Method

Stimulus material, subjects and procedure. Six subjects who had proven unable to code the six-item stimuli efficiently in Experiments III and IV and who were able to demonstrate the ability to whistle a tune, were used.

The six-item stimuli on Tapes F, FAc and FAuc were presented over a loudspeaker in a balanced order. Subjects were requested to "shadow" each stimulus by whistling the "same" notes immediately they were presented over the loudspeaker. The whistles were recorded, together with the background stimuli to which the subjects were responding; the latter were recorded at a lower level than the whistles.

Although it was originally intended that a physical device be used to analyse the frequencies of the whistles, this was found to be unnecessary, since two observers (of probably greater pitch-matching accuracy than a physical system) made reliable judgements of whether or not the pitches of the whistle agreed with those of the background stimulus.

Results and discussion

The number of occasions in which the pitch of the whistle disagreed, even marginally, with that of the stimulus were recorded. Five of the six subjects showed no errors, but ten errors were recorded for the sixth subject. Eight of these errors occurred under FAc conditions. The deviant subject claimed to be aware of his errors and complained that the one second duration of each note was too brief a time to allow him to "home" in on the pitch.

These results clearly discredit hypothesis (*d*), that perceptual rather than storage factors were operating in the previous experiments, since no difficulty at all was experienced in the immediate reproduction of the stimuli under any of the conditions.

General Discussion

The following conclusions may be drawn from the results of the five experiments.

(1) The existence of amplitude as a correlated dimension improved the short-term memory for pitch.

(2) Conversely, correlated frequency improved the retention of intensity.

(3) Of the possible explanations of the phenomenon, all but the influence of informational redundancy upon memory were rejected by experiment.

(4) Correlated duration did not improve the retention of pitch or duration.

(5) Uncorrelated variations in amplitude had an adverse effect upon the retention of pitch.

(6) Knowledge of the form of the relationship between the two dimensions enhanced the improvement shown in the correlated conditions, but probably had no effect upon the deterioration associated with zero correlation between dimensions.

(7) Musically trained subjects transmitted more information than comparable non-musical subjects and were quite uninfluenced by changes between conditions.

In the correlated conditions the memory store contained duplicated information; thus if the primary source failed, the correlated information may still have been available (much as in Baddeley's study the resolution of Z from B may have been aided by the phonemes /ɛ/ and /d/ if the resolution of the initial phonemes failed). The correlated information could therefore serve either to fill in the gaps in memory or as a check upon the chosen code. In the early experiments in which the relation between dimensions was not made known, the improvement under the correlated condition, although significant, was small. With knowledge of the inter-dimensional relationship, performance was greatly improved, suggesting that perhaps only the more sophisticated subjects benefited by the correlated dimension in the earlier experiments.

Conclusion (5), however, introduces another factor whose effect is most forcefully illustrated in Experiment IV. In this experiment, although subjects were warned that there was no relation between pitch and intensity and, in fact, instructed to disregard the intensities, a deterioration in performance still resulted. This implies that the amplitude information was not filtered out at the periphery (Broadbent, 1958), but entered the store and was erroneously used as a source of information. Thus an incomplete memory for pitch order was supplemented by an intensity order which was known to be uninformative. These results are perhaps best explained by the theory that qualities in the stimulus are selected as "filing tags" in memory (Yntema and Trask, 1963; Broadbent, 1967) but adds the rider that this tendency may be strong enough to withstand instructions to ignore the wrong tag.

The results of the musically trained subjects, unlike their controls, suggest that they made use of a peripheral filter. This conclusion is not based upon the fact that their performance did not deteriorate in the uncorrelated condition since they might well have had the capacity to ignore amplitude information which passed through the filter. Rather the conclusion is drawn from the inability of these subjects to make use of the *useful* (correlated) information. It is significant perhaps that a course in musical appreciation consists partly of training the student to widen and narrow his attentive focus at will so that he may listen to one instrument or to many simultaneously. Thus the apparent absence of the peripheral filter in the untrained subject may simply indicate his lack of training in its efficient deployment.

The final outstanding problem concerns the ineffectiveness of duration as a correlated dimension. Firstly it is unlikely that the perception of pitch could have been much affected by the short durations of some of the tones. Although there is evidence that pitch perception is poor at short durations (Turnbull, 1944; Doughty and Garner, 1948), the shortest stimuli were well above those shown to have effects upon the perception of pitch. Also, the variation in duration would be more likely to enhance the phenomenological differences in intensity than reduce them. The possibility that the differences in duration were *not perceptible* can be rejected by some unpublished results of a group of Naval ratings who showed an equal ability to *code* duration as to code intensity and pitch (unpublished data). A clue may be gained from the report by the greater majority of subjects that differences in duration were simply not noticed. This suggests that *changes* from the end of one tone to the beginning of the next were the principal cues used in coding the stimulus.

On the whole the experiments have shown that the kind of parameters which affect search and absolute identification may also affect memory and indicate that discrimination between the features of a memory may be the same kind of function as that between material presented to the senses.

The writers are indebted to the Royal Navy for the use of subjects and equipment, and the U.S. Navy Electronics Laboratory for supplying equipment and the opportunity for preparing the work for publication.

References

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quart. J. exp. Psychol.*, **18**, 302–9.

Broadbent, D. E. (1958). *Perception and Communication*. Oxford: Pergamon.

Broadbent, D. E. (1967). The well ordered mind. *Am. Ed. Res. J.* (in press).

Conrad, R. (1964). Acoustic confusions in immediate memory. *Brit. J. Psychol.*, **55**, 75–84.

Crossman, E. R. F. W. (1960). Information and serial order in human immediate memory. In Cherry, C. (Ed.), *Information Theory*. London: Butterworth.

Doughty, J. M., and Garner, W. R. (1948). Pitch characteristics of short tones. I. Pitch as a function of tonal duration. *J. exp. Psychol.*, **38**, 478–94.

Eriksen, C. W. (1952). Location of objects in a visual display as a function of the number of dimensions on which the objects differ. *J. exp. Psychol.*, **44**, 56–60.

Eriksen, C. W. (1953). Object location in a complex perceptual field. *J. exp. Psychol.*, **45**, 126–32.

Eriksen, C. W., and Hake, H. W. (1955). Multidimensional stimulus differences and accuracy of discrimination. *J. exp. Psychol.*, **50**, 153–60.

Garner, W. R. (1962). *Uncertainty and Structure as Psychological Concepts*. New York: Wiley.

Garner, W. R., and Creelman, C. D. (1964). Effects of redundancy and duration on absolute judgements of visual stimuli. *J. exp. Psychol.*, **67**, 168–72.

Green, B. F., and Anderson, L. K. (1956). Color coding in a visual research task. *J. exp. Psychol.*, **51**, 19–24.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes–loudness. *Amer. J. Psychol.*, **69**, 1–25.

Turnbull, W. M. (1944). Pitch discrimination as a function of tonal duration. *J. exp. Psychol.*, **34**, 302–16.

Yntema, D. B., and Trask, F. P. (1963). Recall as a search process. *J. verb. Learn. verb. Behav.*, **2**, 65–74.

Manuscript received 18th May, 1967.

SOME EFFECTS OF DISCRIMINABLE GOAL-BOX CONDITIONS ON THE LEARNING OF A SUCCESSIVE DISCRIMINATION

BY

M. H. SHELDON*

From the Psychological Laboratory, Cambridge

In an attempt to confirm and extend a previous result, rats were trained on two tasks where a signal delivered at the start of each trial indicated which of two paths through a maze would be rewarded. In Experiment I both paths led to the same goal-box, and it was found that performance was better when the state of the goal-box was different on trials with each of the two signals. In Experiment II the two paths led to spatially separated goal-boxes. It was found that when the states of the two goal-boxes were discriminably different but the state of each of them remained the same from trial to trial, performance was better than when their states varied irregularly. It is suggested that these results have interesting implications for theories of behaviour.

Introduction

In one of the experiments reported briefly in a previous paper (Sheldon, 1964) rats were trained to take a path leading upwards from a starting box when one signal was presented to them, and a path leading downwards from the same starting box when a different signal was presented. For half the animals (Group SGB) the two paths led to a common goal-box, and for half (Group DGB) they led to different goal-boxes. Animals in Group DGB learned the problem significantly faster.

At least two ways of explaining these results in terms of existing theories ought to be considered. The first of these arises as a necessary consequence of the design of the experiment. As far as possible the responses required of the two groups (up or down through the central choice section of the maze) were the same, but it was obviously not possible for them to be identical: for Group SGB both responses took the animal to the same position in space, but for Group DGB they necessarily led to different positions. It might be argued, therefore, that the sequences of movement demanded of the animals trained with a single goal-box were in some way harder to learn than those for the animals trained with two goal-boxes, and that this was the reason for the difference between the groups.

A second explanation relies on secondary reinforcement. In the simple case where animals are trained to choose one path rather than another (but always the same one) we should expect there to be a difference between training with one goal-box and training with two. For animals with only a single goal-box, wrong choices would be secondarily reinforced (since they would lead the animal to a goal-box associated with food) and this would slow down the rate of learning. Animals trained with two goal-boxes would have the advantage that wrong choices would take them to a goal-box that had never been associated with reward. But in a successive discrimination problem (where the two routes are correct equally often) this advantage is lost. For animals in Group DGB both goal-boxes will have been associated with reward, and this means that wrong choices will receive some amount of secondary reinforcement. The critical question, therefore, becomes: do animals trained with two goal-boxes receive *as much* secondary reinforcement for wrong choices as animals

* Now at Department of Psychology, University of Hull.

trained with a single goal-box? One variable that presumably affects the strength of secondary reinforcement is the relationship between the total number of occasions on which the animal has encountered a particular stimulus, and the number of occasions on which it has been rewarded in its presence. As far as this relationship is concerned, the secondary reinforcing strength of the two goal-boxes for Group DGB will be equal and the same as that of the single goal-box for Group SGB: if we take the case of an animal choosing at random, reward will be presented in each goal-box on approximately half the occasions it is visited. But there is also reason to believe that secondary reinforcing strength increases with the *absolute* number of occasions on which a stimulus has been associated with reward. In the present case this must obviously be greater for the single goal-box for Group SGB than for the two goal-boxes for Group DGB: if we again consider the case of an animal choosing at random, the single goal-box will be associated with reward on half the total number of trials, and each of the two goal-boxes on only a quarter of the total. Thus, animals trained with a single goal-box may receive stronger secondary reinforcement for wrong choices, and it could be argued that this is sufficient to explain their inferiority.

The experiments described here look more closely at the effect of distinctive goal-boxes on successive discrimination under conditions to which neither of these explanations could apply.

Experiment I

In the experiment reported previously, and discussed above, animals for whom the two paths led to different goal-boxes learned faster than animals for whom they led to the same goal-box. One way of putting the difference in training conditions for these two groups is to say that for group DGB (but not for Group SGB) reward was consistently associated with different sets of goal-box stimuli on trials with each of the signals. It might be, therefore, that using a single goal-box for all animals, but arranging matters so that on any particular trial it could be in either of two distinctive states, we should find that animals for whom there was a consistent relationship between the signals and the states of the goal-box (and hence between the signals and the stimuli with which reward was associated) learned better than animals for whom there was no such relationship. This experiment tests this possibility.

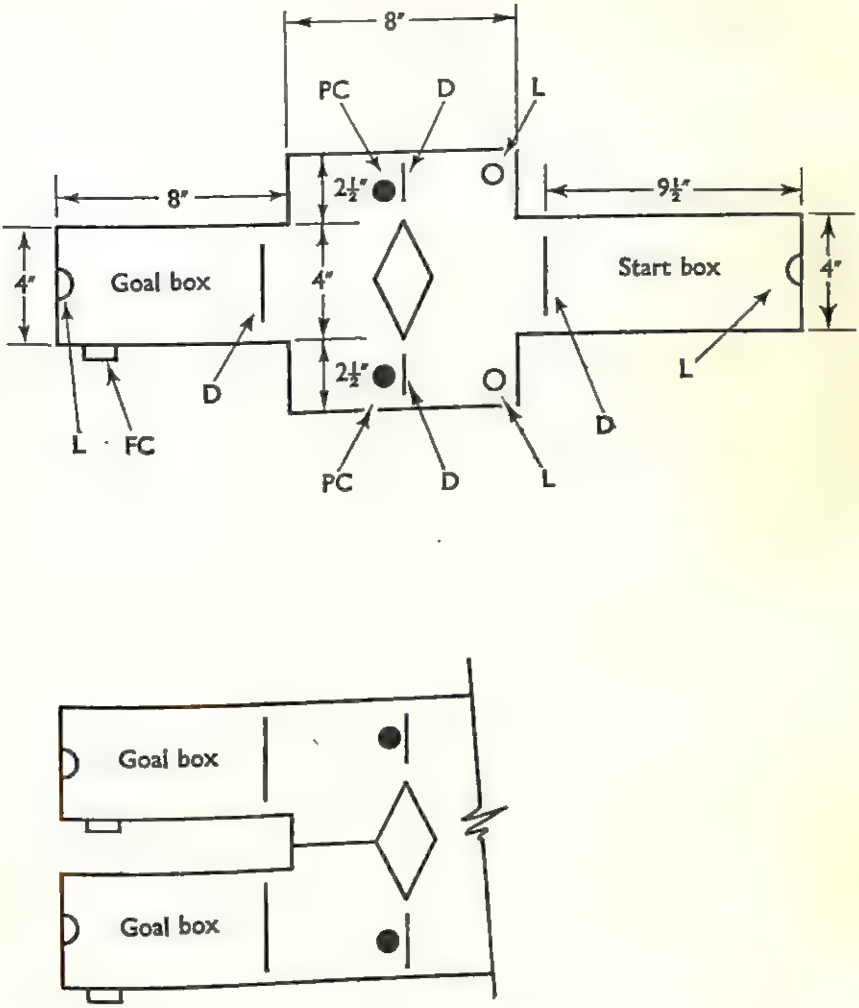
Subjects and apparatus

The subjects were 24 hooded rats (12 males and 12 females) bred in the Psychological Laboratory, Cambridge. At the start of the experiment they were about 90 days old.

The apparatus consisted essentially of a starting box from which pathways led up and down through a central choice compartment to a common goal-box. The upper part of Figure 1 shows a side view of the maze with some of the more important dimensions noted. (It is identical to the single goal-box version of the apparatus described previously (Sheldon, 1964).) The maze was made of unpainted alloy and was lighted at the points indicated in Figure 1 by 4·5 volt bulbs. The depth from front to back (from the rat's point of view the width of the alleys) was 4 in. The front of the maze (the surface between rat and experimenter) was two thicknesses of "Perspex" with a dark filter between them. The lid of the starting box, and the front surface of the goal-box were hinged so that rats could be put into the maze and taken out of it. The operation of all the doors marked in Figure 1 (at the entrance to the starting box, above and below the central obstacle, and at the entrance to the goal-box) was controlled electro-mechanically: the doors were pushed in and out from the back of the apparatus by the arms of solenoids operated by switches. The goal-box contained two food cups side by side; access to them was through holes of 1 in. diameter covered by sliding lids controlled by the arms of separate solenoids.

The apparatus was equipped so that distinctive signals ("starting signals") could be delivered to the rat in the starting box. One signal was the sounding of a buzzer immediately behind the starting box. The other was the flashing of the light in the starting

FIGURE 1



The upper part of the Figure shows the maze used in Experiment I. The maze for Experiment II was similar except that the goal-box unit was as shown in the lower part of the Figure. Key: FC = food cup; PC = photocell; D = door; L = light.

box and the two lights in the choice section of the maze at a rate of about one flash per sec.; the flashing was accompanied by the click of a relay. On trials when the buzzer was sounded none of these three lights was ever on. The signals were switched on manually at the start of each trial, and were switched off by the animal as it passed through the beam of either of the photocells (shown in Figure 1 above and below the central obstacle).

Two interchangeable metal linings were made to fit the goal-box. Both linings covered the end wall of the goal-box (leaving a hole for the light bulb to stick through) and extended for $5\frac{1}{2}$ in. along the floor with a hole giving access to one of the food cups. One insert was painted black, and had a rough floor (small nuts and bolts set in a layer of glue); a vertical baffle attached to the left-hand side of the surface covering the end wall of the goal-box forced the animal to eat from the right hand food-cup. The other lining was painted white, and had a smooth floor; a baffle at the right-hand side forced the animal to eat from the left-hand food cup.

Pretraining

Some days before pretraining started animals were placed on a 23-hr. food deprivation schedule, and this was maintained throughout the experiment. Food was available in the home cage for 1 hr. each day, shortly after the end of the experimental session. On each of the 4 days of pretraining animals were placed directly into the goal-box and allowed to eat freely. Both linings were in place equally often.

Training and design

After pretraining animals were divided randomly into two groups of 12, with the restriction that each group contained six males and six females. One group (which we will call Group S) was trained with a systematic relationship between starting signals and goal-box linings. For six animals the goal-box lining was always white on trials with the buzzer, and black on trials with the flashing light; for the remaining animals this was reversed. Within each of these sub-groups half the animals were trained to go up for the buzzer and down for the light, and half to go down for the buzzer and up for the light. The other group (Group R) was trained with identical starting signals and goal-box linings but with the relationship between them random. Equal numbers of animals were trained in each direction to each starting signal.

Both groups had six trials a day. The starting signals were delivered according to a selected sequence repeating every sixth day, and ensuring that there were three trials with each signal on each day. For Group R a similar sequence determined the goal-box lining on each trial; it ensured that each goal-box lining was paired equally often with each signal. Animals were tested in two batches of 12: all the males had all their trials first, and then all the females. Each batch thus contained six animals from each group. Within each batch of animals trials were given in rotation: between successive trials an animal was kept in an individual cage in the testing room while a trial was given to the other members of the batch. The interval between trials was about 10 min.

The details of trial-by-trial procedure were as follows. Once the appropriate lining had been fitted in the goal-box the trial started with the rat's being placed in the starting box. Five sec. later the starting signal was switched on, and 3 sec. after that the door of the starting box was opened leaving the rat free to make its choice. The signal stayed on until the rat passed through the beam of one of the photocells. As soon as the rat entered the goal-box the door was closed behind it. If it had chosen correctly it was allowed to eat for 4 sec. from the food cup (which contained small pieces of normal laboratory diet). If it had chosen wrongly it was kept in the empty goal-box for 4 sec. After wrong choices a correction procedure was used. The rat was returned to the starting box and allowed a second free trial, under conditions identical to those on the first trial. If this was also wrong it was given a third trial on which it was *forced* (by closing the door either over or under the central obstacle) to take the correct path.

Training went on for 38 days, until all the animals in one group had reached the criterion of at least five correct trials out of six on four successive days.

Results and discussion

By the time the last animal in Group S had reached the criterion only nine animals in Group R had reached it. If we assume that from this point animals who had not

TABLE I

MEAN NUMBER OF TRIALS CORRECT OVER THE WHOLE TRAINING PERIOD AND OVER EACH HALF TAKEN SEPARATELY. (EXPERIMENT I)

Trials	*Group S*	*Group R*	*p**
1–228	176·5	156·7	<0·025
1–114	77·5	71·0	>0·05
115–228	98·9	85·6	<0·01

* Probabilities are one-tailed as calculated by the Mann-Whitney U test.

yet reached the criterion would have performed perfectly if the experiment had continued, we can compare members of the two groups for the number of trials to reach criterion. The median numbers of trials were: Group S, 78; Group R, 192. By the Mann-Whitney U test, $U = 24 \cdot 5$, $p < 0 \cdot 01$, one-tailed. Table I presents a comparison between groups for the mean number of trials correct over the whole training period, and over each half of it. Taking the data from correction trials and comparing the groups with respect to the percentages of incorrect first choices that were followed by repetitions of the same error (and then, of course, by a forced trial) we find means as follows: Group S, 57·3 per cent.; Group R, 48·2 per cent. Members of Group S, that is to say, were more likely to make a wrong second choice than members of Group R ($U = 36 \cdot 5$, $p < 0 \cdot 025$, one-tailed).

The performance of Group S (for whom there was a systematic relationship between starting signals and goal-box linings) was better, in terms of trials to criterion and initial correct choices, than that of Group R. The differences between groups are similar to those between Group DGB and Group SGB in the experiment reported previously. There is one exception to this that may be important. An unpublished analysis of the results from correction trials in this previous experiment shows that the group that learned the discrimination faster (Group DGB) also made repetitive errors on a smaller percentage of their incorrect trials ($p < 0 \cdot 001$). In the present experiment the group that learned faster (Group S) made *more* repetitive errors.

In the Introduction we considered two possible explanations of the previous results. Neither of them could apply to the present case. Here, with a single goal-box for all animals, responses for the two groups were strictly identical. The secondary reinforcement for wrong choices must also be the same for both groups: both the relative and absolute frequencies of reward associated with the two states of the goal-box will be the same.

Experiment II

In Experiment I we found that there was an advantage in manipulating the state of a single goal-box in such a way that reward was consistently associated with a different set of stimuli on trials with each of the two starting signals. The next experiment tried to produce analogous results with two goal-boxes.

Two different and interchangeable goal-box linings were used. For one group of animals (Group S) the linings stayed in the same positions (one in each goal-box) throughout the experiment. For the other group (Group R) the same linings were used, but their positions shifted randomly between trials. Both groups had to learn a successive discrimination. For Group S reward on trials with a particular starting signal would always occur in the presence of the same goal-box lining. For Group R no such relationship would exist. It was predicted (by analogy with the previous experiments) that Group S would learn the problem faster.

Subjects and apparatus

The subjects were 22 hooded female rats, about 90 days old at the start of the experiment. The apparatus was the same as that used in Experiment I except that a second goal-box was added, and a temporary floor was fitted behind the rear edge of the central obstacle to separate the upper and lower paths. These changes are shown in the lower part of Figure 1. The second goal-box was designed to be as similar as possible to the first. In all other ways the maze was left unchanged.

The linings for the goal-boxes were modified to make them fill the goal-boxes even more completely. They were lengthened so that they extended 7 in. along the floor, and side pieces were fitted to them to cover the back walls of the goal-boxes. In other respects they were the same as those used in Experiment I.

Pretraining

Animals were placed on a 23-hr. deprivation schedule, and then pretrained for a total of 9 days. On each of the first 3 days they were allowed to eat freely in plain goal-boxes. On the next 3 days the linings were introduced, and each rat fed once each day in the presence of each of them. On the last 3 days animals were given a series of forced trials through the maze: no starting signals were used, and each animal was forced once in each direction on each day.

Design and training

After pretraining the animals were divided into two groups of 11. For Group S the linings were always in the same position. For six animals the upper goal-box always contained the black insert and the lower one the white; for five animals this relationship was reversed. For Group R the position of the linings varied from trial to trial according to a schedule that satisfied the following conditions: (*a*) each spatial arrangement of the linings (black above—white below, and white above—black below) occurred equally often; (*b*) each of these arrangements was paired equally often with each starting signal; (*c*) reward was associated with the two inserts equally often.

The animals were trained in two batches of 11 (run at different times of day) each batch containing six animals from one group and five from the other. In all other respects the training procedure was the same as that in Experiment I.

Training continued for 32 days, until all the animals in one group had reached the criterion of at least five correct trials out of six on four successive days. One animal in Group S was removed from the experiment on day 19; it refused to run in the apparatus although it had successfully met the criterion some time previously. Its data are included where appropriate.

Results and discussion

By the time the last animal in Group S had reached the criterion only six animals in Group R had reached it. Assuming that from this point animals who had not reached criterion would have performed perfectly if the experiment had been continued, we can compare the two groups for the number of trials to criterion. The median numbers of trials were: Group S, 144; Group R, 174 ($U = 31{\cdot}5$, $p < 0{\cdot}05$, one-tailed). Table II presents a comparison between groups for the mean number of trials correct over the whole training period and over each half separately. Taking the data from correction trials, and comparing the groups with respect to the percentages of incorrect first choices that were followed by repetitions of the same error (and then, of course, by a forced trial) we find mean scores as follows: Group S, 12·6 per cent.; Group R, 20·6 per cent. Members of Group S, that is to say, were significantly less likely than members of Group R to make a wrong second choice ($U = 19$, $p < 0{\cdot}01$, one-tailed).

TABLE II

Mean Number of Trials Correct Over the Whole Training Period and Over Each Half Taken Separately. (Experiment II)

Trials	*Group S*	*Group R*	p*
1–192	141·0	122·4	<0·025
1–96	59·5	55·3	>0·05
97–192	81·5	67·0	<0·01

* Probabilities are one-tailed as calculated by the Mann-Whitney U test.

In this experiment, as in Experiment I, the responses, and the secondary reinforcement for wrong choices, must be considered the same for both groups. Nevertheless, rats in Group S reached the criterion significantly more quickly, and made fewer

errors, particularly during the second half of training. These differences between groups are similar to those in Experiment I. However, in Experiment II the group that learned more quickly also made fewer errors on the second choices of correction trials. In this the results resemble those of a previous experiment (Sheldon, 1964) but not those of Experiment I.

General Discussion

From this series of experiments it appears that under certain circumstances rats find it easier to learn successive discriminations when reward occurs in the presence of different sets of stimuli with each of the two starting signals. In the experiment reported previously the distinctive stimuli were those provided by two physically distinct goal-boxes. In the experiments reported here they were presumed to be stimuli arising from the interchangeable linings of the goal-boxes.

These results present certain theoretical difficulties. Theories of behaviour have usually allowed for only one kind of relationship between stimulus and response. For S-R theory—even its more sophisticated versions—the presentation of a stimulus is the occasion for the animal to make a certain set of movements. For other theories —Spence's theory of discrimination learning (1936), Deutsch's theory (1953, 1960) and the most recent version of Mowrer's theory (1960)—the relationship is quite different: the function of a stimulus is to serve as a cue for the animal to approach. Neither of these functions alone seems sufficient to explain the present results.

For S-R theory the central difficulty is that since the stimulus conditions up to the moment of choice, and the primary and secondary reinforcement contingent upon the choice, were apparently identical, there is no basis for predicting differences between the groups in either experiment. For "approach" theories there is a different difficulty. In these experiments behaviour cannot be regarded as consisting *simply* of approaches to sets of goal-box stimuli, since which set of stimuli is to be approached on a particular trial depends on which starting signal is presented. But these theories as they stand have difficulty in accommodating the function of the starting signals at all, since they are stimuli that the animal is not in any sense required to approach. In terms of Deutsch's theory, for example, each of the starting signals would have to be represented by a link at the bottom of the chains of links representing the routes through the maze. As soon as the signal was presented to the animal in the starting box its link would be turned off before the signal could have played any part in selecting the animal's response. There are similar difficulties for Mowrer's theory. Behaviour in many situations is considered to consist of approaches to stimuli (exteroceptive or proprioceptive) to which certain patterns of autonomic response have been conditioned; again no provision seems to exist for behaviour to be controlled by stimuli like the starting signals.

What seems to be required is some theoretical arrangement by which certain stimuli (like the starting signals in these experiments) can serve to *select* other stimuli (like the stimuli arising from the goal-boxes) as the goals for approach responses on a particular trial. A starting signal would select as goal any set of stimuli that—on trials using that signal—was reliably associated with reward; once that set of goal stimuli had been selected, the animal's task would be to produce a suitable approach response. This kind of arrangement would explain the faster learning of the "systematic" groups in these experiments: for them (but not for the "random" groups) reward was given in the presence of different sets of stimuli on trials with the two starting signals. Each starting signal, therefore, would be able to select a different set of stimuli to be approached. As it stands this explanation has two weaknesses. In the

first place it predicts that the "random" groups (for whom the starting signals would not be able to select different sets of goal-box stimuli) should have found learning impossible and not merely difficult. In the second place, it needs to specify far more clearly how it comes about, given that a particular set of goal stimuli has been selected, that an "approach" response is organized.

It should be pointed out that in two further sets of experiments no effect of discriminable goal-box conditions was found. One set involved successive discrimination in a jumping stand, and the other successive discrimination between left and right in a horizontal maze using shock avoidance as reinforcement. These two experimental situations were intended to present rats with problems formally analogous to those described here. On the other hand, the present findings agree well in some respects with those of an experiment by Shepp (1962) in which the subjects were mentally retarded children. Shepp found that a successive discrimination (responding to the left for one pair of objects and to the right for another) was easier when the two responses had different rewards associated with them. The rewards can be seen as fulfilling the same functions as the discriminable goal-box stimuli in the present experiments.

I should like to thank the editorial referee who drew my attention to the experiment by Shepp. The experiments reported here were done at the Psychological Laboratory, Cambridge, while I held a Scholarship for Training in Research Methods from the Medical Research Council. It is a pleasure to acknowledge the advice of Mr. A. J. Watson.

References

Deutsch, J. A. (1953). A new type of behaviour theory. *Brit. J. Psychol.*, **44**, 304–17.
Deutsch, J. A. (1960). *The Structural Basis of Behaviour.* London: Cambridge University Press.
Mowrer, O. H. (1960). *Learning Theory and Behaviour.* New York: Wiley.
Sheldon, M. H. (1964). Response learning. *Nature*, **202**, 1141–2.
Shepp, B. E. (1962). Some cue properties of anticipated rewards in discrimination learning of retardates. *J. comp. physiol. Psychol.*, **55**, 856–9.
Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychol. Rev.* **43**, 427–49.

Manuscript received 20th June, 1967.

THE BINOMIAL DISTRIBUTION OF RIGHT, MIXED AND LEFT HANDEDNESS

BY

MARIAN ANNETT

From the Department of Psychology, University of Hull

Right, mixed and left handers are found in binomial proportions in seven samples of varied subjects whose lateral preferences were ascertained by several methods. These proportions have been obtained in previous studies of humans and animals when the performance of several actions has been recorded in complete samples and when consistent right and left subjects have been separated from those of mixed usage.

INTRODUCTION

The classification of varieties of handedness is a problem basic to all enquiry into lateral differences, their inheritance and their implications for the localization of function in the cerebral hemispheres. The simplest classification recognizes only two types of handedness, right and left. The fact that the majority of self-styled left handers use the right for some skilled unimanual acts is often attributed to the effects of diverse cultural pressures on the natural sinistral to use the right hand. On this view, left preference for any skilled one handed action is taken as evidence of basic sinistrality and warrants classification as a left hander. These assumptions seem to have governed the majority of studies of the genetic and physiological concomitants of left handedness and the "left" handers reported in the literature are a heterogeneous collection of individuals varying in degree and consistency of preference for the left hand.

A second classification recognizes three varieties of manual preference, right, left and mixed; the last usually involves a preference for the right hand in some skilled acts and the left for others but would include also the rare cases of indiscriminate use of either hand. This threefold classification has been made most often in work with children with developmental dysphasias and dyslexias (Orton, 1937; Ingram, 1960) and with the notable exception of Chesher (1936) has not been used until recently in descriptions of patients with neurological disabilities. The possibility that the mixed hander is a "basic biological variant" and not simply a shifted sinistral has been suggested by Gillies, MacSweeny and Zangwill (1960) who described left handed writers preferring the right hand for actions presumably less subject than writing to cultural pressures. If this is the case, many of the problems concerning the inheritance and cerebral dominance of left handers might have arisen from a failure to distinguish mixed from pure left cases. A classic Mendelian model of the genetic basis of handedness and hemispheric dominance has been used to suggest how mixed handers might respond differently from left handers to similar cerebral insults (Annett, 1964). Pure right handers were supposed to be dominant homozygotes (having two "genes" for right handedness) pure left handers recessive homozygotes (having two for left) and mixed handers heterozygotes (having one "gene" for left and one for right handedness).

A third view of the classification of handedness might be that all groupings, whether into two, three or more categories, entail arbitrary divisions of a continuous distribution of lateral differences. On the basis of Galton's data on strength of grip, Woo and

Pearson (1927) concluded that, "Dextrality and sinistrality are not opposed alternatives but quantities capable of taking values of continuous intensity and passing one into the other." Differences between the hands formed a bell-shaped distribution whose mode showed the advantage of the right hand, but "the ordinate of this curve at zero difference occupies its proper place in the continuous series of ordinates" (Woo, 1928). Studies of handedness more often concern preference than measures of strength or skill. Distributions of degree of preference have been derived by several systems of scoring the numbers of actions performed with each or both hands (Hildreth, 1949; Naidoo, 1961; Crovitz and Zener, 1962). These distributions usually take the form of a J with the anti-mode at or near the point of no difference, or equal preference for both hands. Although the distributions for strength and preference take such different forms, both appear to be continuous. Any classification of varieties of lateral asymmetry is therefore to some extent arbitrary. The purpose of the model referred to above was to suggest that some advantage might be gained from a novel treatment of the distribution. Instead of single division at a variable point in the middle, it was proposed to cut off the two extremes, the pure right and pure left handers, and to treat all remaining cases as mixed handers. A monofactorial model of inheritance was used because it was sufficient to illustrate the possible advantages of separating mixed handers from consistent ones. The genetic basis of handedness may well be more complex, but this does not affect the argument that mixed handers might be more apt than consistent ones to develop the skills of the alternate hand and cerebral hemisphere.

The collection of the data reported below was begun with the aim of discovering whether the consistent left handers, postulated by the model, exist. Humphrey (1951) has demonstrated the inconsistency of left preference of those who regard themselves as left handed. It seemed possible that cultural pressures might force all natural sinistrals to acquire at least some right handed habits. It was important, therefore, to see whether individuals could be found who are as consistently left handed as most right handers are consistently right, and in what proportions they might occur.

After collecting the first few samples, it was noticed that the three varieties of handedness were occurring in binomial proportions. These proportions would be predicted by the simple Mendelian model outlined above if the factor for left handedness were always expressed in heterozygotes. (The prediction had not, in fact, been made for it was assumed that only a few heterozygotes in normal populations would reveal their left tendencies.) It was then considered important to check the reliability of the finding of binomial proportions with improved methods of sampling, subjects other than university students and methods of enquiry other than questionnaire. It is the object of this paper to report this one most interesting aspect of the data.

Method

A. *Questionnaire* 1

Among questions of opinion as to own and family handedness, the first questionnaire asked which hand was used to perform eight actions. (These were: (*a*) to write a letter legibly, (*b*) to throw a ball to hit a target, (*c*) to play a game requiring the use of a racket such as tennis, (*d*) at the top of a broom to sweep dust from the floor, (*e*) at the top of a shovel to move sand, (*f*) to hold a match when striking it, (*g*) to hold scissors to cut paper, (*h*) to hold a thread to guide through the eye of a needle (or guide a needle onto a thread).

Individuals were classified into three groups as follows:

(i) *right handers* who use the left hand for none of the actions;

(ii) *mixed handers* who report a mixture of right and left preference; and

(iii) *left handers* who use the right for none of the eight actions.

("Either" responses were not sufficient for mixed classification; right plus either responses were counted right and left plus either counted left).

Data were obtained for three groups of university students, described in Table I, A. In each group, questionnaires were given out and requested to be handed in completed at a later date. All samples and particularly the last which was collected through the university post, fell short of 100 per cent. returns.

Twins were excluded from these and subsequent samples in view of reports that twins are more likely to be left handed than the single born. Non-British students were omitted from University of Hull samples. (Information as to nationality was not obtained for other samples.)

B. *Questionnaire 2*

A second questionnaire was prepared with the aim of giving as full an opportunity as possible for mixed handers to be differentiated from consistent ones. Four more questions were added to the eight used in the first questionnaire: (*a*) to deal playing cards, (*b*) to hammer a nail into wood, (*c*) to hold a tooth-brush while cleaning the teeth, (*d*) to unscrew the lid of a jar. Respondents answering consistently right or left for all actions were asked what other skilled, unimanual actions, if any, would be performed with the alternate hand.

Individuals were classified as before. Acts said to be done by the otherwise non-preferred hand were in most cases ones usually done with that hand such as changing car gears, smoking, using a fork; two actions regarded as sufficient to place subjects in the mixed category were playing darts and wielding a cricket bat left handed.

The two groups responding to this questionnaire, described in Table I, B, were complete samples since the forms were completed and returned immediately in class. The students are different individuals from those completing the first questionnaire, except for a few repeating courses. The second sample is of a non-university population.

C. *Observed handedness*

Observations of hand usage have been made for two samples described in Table I, C. The school children were observed performing five actions: (*a*) pointing (to indicate choice responses), (*b*) drawing, (*c*) throwing a ball, (*d*) cutting with scissors and (*e*) unscrewing the lid of a jar.

The students taking courses in psychology, observed one another during laboratory classes performing the 12 actions used in the questionnaire. Most of the students had filled in one of the two questionnaires at least 3 months before the class.

TABLE I

INCIDENCE OF RIGHT, MIXED AND LEFT HANDEDNESS IN SEVEN SAMPLES

		Right		*Mixed*		*Left*		*Total*	*Chi square 1 d.f.*
		N.	*per cent.*	*N.*	*per cent.*	*N.*	*per cent.*		
A. *Questionnaire 1*									
1. Aberdeen First year Psychology Students (85 per cent. returns)	N. observed	256	(71·5)	90	(25·1)	12	(3·4)	358	1·400
	N. expected	253·0		96·0		9·0			
2. Hull Psychology Students (96 per cent. returns)	N. observed	175	(67·3)	72	(27·7)	13	(5·0)	260	2·436
	N. expected	171·2		79·7		9·16			
3. Hull sample of all students (54 per cent. returns)	N. observed	80	(71·4)	29	(25·9)	3	(2·7)	112	0·057
	N. expected	79·7		29·7		2·67			
B. *Questionnaire 2*									
4. Hull First Year Psychology Students	N. observed	99	(70·7)	36	(25·7)	5	(3·6)	140	0·638
	N. expected	97·7		38·6		3·71			
5. Enlisted Men	N. observed	85	(63·9)	43	(32·3)	5	(3·8)	133	0·012
	N. expected	85·2		42·6		5·2			
C. *Demonstration*									
6. Schoolchildren 5–15 years City of Hull	N. observed	101	(58·4)	64	(37·0)	8	(4·6)	173	0·247
	N. expected	102·2		61·7		9·2			
7. Hull Honours Psychology Students	N. observed	31	(62·0)	18	(36·0)	1	(2·0)	50	0·007
	N. expected	31·9		16·2		1·9			

Results

The results given in Table I include both males and females since a comparison of the sexes showed no statistically significant differences; the proportions of males in the pure left group was slightly greater than that of females.

The table gives the numbers of individuals in each sample who performed all actions consistently with the right hand, all consistently with the left hand and those showing any mixture of usage. The percentages of those called left are smaller and those called mixed are larger than in many previous studies; this follows from the criteria of classification adopted for reasons described in the introduction.

The chief interest of the table lies in the fact that in every sample, the observed numbers agree closely with those which would be expected on the assumption of a simple monofactorial model of the genetic basis of handedness; that is, if right handers have two factors for right, left handers two for left and mixed handers one of each type, the factors being combined in the proportions $r^2{:}2rl{:}l^2$. (For reasons to be mentioned later this model is in fact insufficient to account for the inheritance of handedness; the point being established here is that the phenotypes occur in these proportions.)

The expected numbers have been calculated according to Levene's formula which is appropriate when one of the factors has a low frequency (Li, 1961, p. 29). The calculations are made as follows, group 1 being used as an example. Let the factor for right be g_1 and the factor for left g_2.

$$g_1 = 2 \text{ right} + \text{mixed} = (2 \times 256) + 90 = 602.$$

$$g_2 = 2 \text{ left} + \text{mixed} = (2 \times 12) + 90 = 114.$$

$$2G = \text{Total factors in the population} = 716.$$

$$\text{N. right expected} = \frac{g_1(g_1-1)}{2(2G-1)} = 253{\cdot}01$$

$$\text{N. mixed expected} = \frac{g_1 g_2}{2G-1} = 95{\cdot}98$$

$$\text{N. left expected} = \frac{g_2(g_2-1)}{2(2G-1)} = 9{\cdot}01$$

In none of the seven samples do the observed numbers differ significantly from those expected and in most samples the fit is extremely close.

Discussion

The first possibility requiring examination is that the binomial proportions found here are an artefact of some aspect of the method of enquiry. It was to test this possibility that several samples were collected using varied methods of investigation (questionnaire and observation), and varied subjects (students, enlisted men and school children). In one sample, the observations were not made by the writer but by students in laboratory class. Another variable was the number of actions involved; the first questionnaire asked eight questions, the second 13 (including an open question) and the children were observed performing only five actions. Although any distribution could be divided into three parts in binomial proportion *after* it was obtained, the chances are very small of obtaining these proportions when the criteria of division are fixed in advance and when different numbers of actions are

involved in different samples. The only condition was that all acts should be performed consistently with one hand, right or left, and that all combinations be regarded as mixed. There seems no reason, therefore, for regarding the binomial proportions as an experimental artefact.

It may be useful to consider briefly some of the conditions without which binomial proportions would be unlikely to occur. One is the inclusion of actions varying widely in the degree of skill demanded of the left hand. Actions such as sweeping and unscrewing the lid of a jar must be used so that those with only mild sinistral tendencies are revealed; at the other extreme, the action of cutting with scissors is required to distinguish those so strongly left handed that they must learn to use ordinary scissors (not special left handed ones) with that hand. Another condition normally required is that complete samples of subjects should be obtained. In a survey of a large number of students whose cooperation could be only invited, the proportion of right handers was too small for binomial proportion; those with left tendencies more often volunteered. (One exception to this rule seems to be group 3 in Table I; this group does not differ from groups 1 and 2 for several characteristics and it has been concluded that sampling factors are unrelated to handedness in this case.) A third condition of the appearance of binomial proportions is that the criteria of classification should be exactly those adopted here. In previous studies where the criteria of right and left handedness have been more flexible, the proportion of mixed handers has been too small.

There are in the literature, however, several reports in which the basic data are given in sufficient detail for reclassification according to the present criteria. Clark's (1957) study of 12 instances of hand usage in Glasgow school children found 236 right, 68 mixed and eight left handers. These numbers are almost identical with those predicted (Chi square 0·006). Binomial proportions may be found also in data collected in the U.S.A. by Harris (1957) and by Merrell (1957). Sutton's (1963) findings for 772 Australians are in these proportions but not those for 257 Polynesians among whom insufficient mixed handers were recorded.

Finch (1941) found for chimpanzees and Warren (1953) for monkeys, that the numbers of animals using the left hand consistently (over 90 per cent. of trials) equalled those using the right. In this case the binomial requires half the animals tested to be mixed handed; ambilaterals were 40 per cent. of 30 chimpanzees and 45 per cent. of 84 monkeys. In two studies of monkeys, Ettlinger (1961, 1964) found more left than right handers but in both groups the ambilaterals outnumbered either; the numbers of animals in each handedness group early in training, but not those at the end of training, follow binomial proportions closely (Chi square less than 1 for both sets of data). Cole's (1955) results for 60 cats also fit binomial expectations.

There seems little room for doubt, therefore, that when limb preferences are ascertained for several actions in complete samples of subjects and when all individuals showing mixed tendencies are separated from those who are consistently right or left, the three varieties occur in binomial proportions. Whatever the factors may be which promote the skilled use of each hand, these factors are randomly combined in the population. Furthermore, it would seem that when both factors are present, both are expressed in at least some tendency to use both hands in unimanual acts. Differences between species in the proportions of individuals showing each type of preference may be due to differences in the proportions of the two factors. Tendencies to use the left limb seem more numerous in animals than humans, but the rules of combination of factors seem to be the same in the several species for which information is available.

There are some other aspects of lateral asymmetry in which the relevant combinations occur at random. Merrell (1957) compared persons who were completely right or completely left handed on five tests for dominant eye. (Merrell did not use the classification "mixed hander" but by subtracting these consistent ones from his original group, the numbers of mixed handers could be deduced for the handedness calculations referred to above.) The right hand and right eye was dominant in 281 (68 per cent.) the left hand and left eye in 11 (2·7 per cent.) and the dominance of eye and hand differed in 121 (29·3 per cent.). These numbers are almost exactly in binomial proportion and the percentages of each combination are remarkably similar to those found for the three varieties of handedness (Table I).

Another example of random combination of right and left, noted by Merrell, is that among twin pairs. If the results of three studies cited by Fuller and Thompson (1960, p. 159) are added together, the numbers of R-R, R-L and L-L pairs are found to occur in binomial proportions for both monozygotic and dizygotic pairs. The proportions of each kind of pair are not unlike the proportions found here for the three kinds of handedness.

The observation of random combination of right and left tendencies in individuals, of right and left handedness in twin pairs and of hand eye dominance should permit some old problems to be approached in a fresh light. The observation does not suggest any immediate solutions but rather opens up new lines of enquiry. Discussions of the respective contributions of nature and nurture to lateral preference have drawn attention to the many kinds of environmental pressure, from social prejudice to the design of tools, which would tend to induce dextral habits (Burt, 1958; Clark, 1957). Such factors may have operated to suppress the left tendencies of some mixed handers in the present series, most of whom would generally be regarded as right handers. These environmental influences are most unlikely, however, to have produced their sinistral inclinations. The criteria of classification used here have the advantage of showing how many individuals have such inclinations; though in most mixed handers the left tendencies are slight, they are nevertheless present in fairly constant proportions from sample to sample. The discovery of pure left handers who are as consistently left as most right handers are consistently right shows that there are some individuals on whom the supposed environmental pressures have had no observable effect. A non-genetic theory of the origin of left handedness is that of Blau (1946) who maintained it to be "an expression of infantile negativism." Even if this were the case, an explanation would still be required of the stable proportions of negativism from sample to sample and especially the combination of negativism and conformity represented by mixed handers. The main problem presented by theories of environmental influence is the difficulty of envisaging how such influences could work to produce an assortment of factors which is so random that it repeatedly and exactly obeys the laws of chance.

A random assortment of factors of similar frequency from sample to sample would seem more likely to be the product of a genetic mechanism than an environmental one. The chief difficulty for theories of the inheritance of handedness is the observation of R-L pairs of monozygotic twins. There is no doubt that inheritance plays some part in the determination of handedness, even of twins (Rife, 1940) but the mechanisms involved remain to be elucidated. There is scope for research in that genetic studies have not distinguished mixed handers from consistent ones. Twins returning questionnaires in the present enquiry include several mixed handers.

The random combination of right and left tendencies found here for one generation would require, in itself, no more that the simple monofactorial model of inheritance

described in the introduction. Genetic theories need to be tested on more than one generation, however, and sufficient families have been collected by the author to show that the simple model cannot account for the facts, even when mixed handers are differentiated in both generations. It would seem that the discovery of binomial proportions under the guidance of a model which would predict them must be attributed to serendipity. Binomial proportions of phenotypes can arise from several polygenic foundations as well as from monogenic ones (Li, 1961, p. 129). Their repeated occurrence in so many studies, both human and animal, encourages the search for a genetic basis for variations in lateral asymmetry and the belief that the separation of mixed handers from consistent ones is an essential first step.

This work was supported by a grant from the Spastics Society. I am grateful to Professor Elizabeth Fraser for facilities at the University of Aberdeen, to Professor A. D. B. Clarke for those at the University of Hull and to Mr. S. W. Hobson, Chief Education Officer to the City of Hull, who kindly permitted the examination of school children. I am indebted to Mr. A. Royse for his interest and advice on the analysis of the data.

References

ANNETT, M. (1964). A model of the inheritance of handedness and cerebral dominance. *Nature*, **204**, 59–60.

BLAU, A. (1946). *The Master Hand*. Research Monograph. No. 5, Amer. Orthopsychiat. Ass.

BURT, C. (1958). *The Backward Child*. London: Univ. of London Press.

CHESHER, E. C. (1936). Some observations concerning the relation of handedness to the language mechanism. *Bull. Neurol. Inst. N.Y.*, **4**, 556–62.

CLARK, M. M. (1957). *Left-Handedness*. London: Univ. of London Press.

COLE, J. (1955). Paw preference in cats related to hand preference in animals and men. *J. comp. physiol. Psychol.*, **48**, 137–40.

CROVITZ, H. F. and ZENER, K. A. (1962). A group test for assessing hand and eye dominance. *Amer. J. Psychol.*, **75**, 271–6.

ETTLINGER, G. (1961). Lateral preferences in monkeys. *Behaviour*, **17**, 275–87.

ETTLINGER, G. (1964). Lateral preferences in the monkey. *Nature*, **204**, 606.

FINCH, G. (1941). Chimpanzees' handedness. *Science*, **94**, 117–8.

FULLER, J. L. and THOMPSON, W. R. (1960). *Behaviour Genetics*. New York: Wiley.

GILLIES, S. M., MACSWEENEY, D. A., and ZANGWILL, O. L. (1960). A note on some unusual handedness patterns. *Quart. J. exp. Psychol.*, **12**, 113–6.

HARRIS, A. J. (1957). Lateral dominance, directional confusion and reading disability. *J. Psychol.*, **44**, 283–94.

HILDRETH, G. (1949). The development and training of hand dominance. *J. genet. Psychol.*, **75**, 221–54.

HUMPHREY, M. E. (1951). Consistency of hand usage. *Brit. J. educ. Psychol.*, **21**, 214–25.

INGRAM, T. T. S. (1960). Paediatric aspects of specific developmental dysphasia, dyslexia and dysgraphia. *Cerebral Palsy Bull.*, **2**, 254–77.

LI, C. C. (1961). *Human Genetics*. New York: McGraw Hill.

MERRELL, D. J. (1957). Dominance of eye and hand. *Human Biol.*, **29**, 314–28.

NAIDOO, S. (1961). *An Investigation into Some Aspects of Ambiguous Handedness*. Unpubl. M.A. Educ. Thesis, Univ. of London.

ORTON, S. T. (1937). *Reading, Writing and Speech Problems in Children*. London: Chapman & Hall.

RIFE, D. C. (1940). Handedness with special reference to twins. *Genetics*, **25**, 178–86.

SUTTON, P. R. N. (1963). Handedness and facial asymmetry: lateral position of the nose in two racial groups. *Nature*, **198**, 909.

WARREN, J. M. (1953). Handedness in the rhesus monkey. *Science*, **118**, 622–3.

WOO, T. L., and PEARSON, K. (1927). Dextrality and sinistrality of hand and eye. *Biometrika*, **19**, 165–99.

WOO, T. L. (1928). Dextrality and sinistrality of hand and eye. Second memoir. *Biometrika*, **22A**, 79–148.

Manuscript received 16th May, 1967.

SHORTER ARTICLES AND NOTES

LEFT-RIGHT DIFFERENCES IN AUDITORY PERCEPTION OF VERBAL AND NON-VERBAL MATERIAL BY CHILDREN

BY

DIRK J. BAKKER

From the Paedological Institute, Amsterdam

One hundred and twenty children, 60 boys and 60 girls, varying in age between 6 and 12 years were presented with a series of digits and Morse-like sound patterns to each ear separately. As predicted, sound patterns were found to be better retained when presented to the left ear than when presented to the right ear. Series of digits however were not retained better via the right ear than via the left ear. The dominance of the left ear for non-verbal material decreases with increasing age. For verbal material a quadratic relation between the dominance of the right ear and age was established.

Introduction

After various research projects with adults it has become apparent that verbal material presented to the right ear is retained better than when presented to the left ear (Kimura, 1961, 1964; Bryden, 1963; Broadbent and Gregory, 1964). On the other hand non-verbal material is retained better when input is via the left ear (Kimura, 1964).

This ear asymmetry has—as far as we know—not yet been investigated in a genetic context, as contrasted with other aspects of lateral dominance, such as hand, eye and foot preference (Gesell and Ames, 1947; Belmont and Birch, 1963; compare Palmer, 1964). These research projects show that an increase in age is accompanied by increasing left or right dominance.

Differences between boys and girls in connection with lateral dominance are generally not found (Belmont and Birch, 1963).

Kimura (1964) mentions that the way in which the material is presented, that is, dichotically or to each ear separately at separate moments, influences results. She found that under the first mentioned condition asymmetry between the ears appears, whilst under the second condition no asymmetry was found. A possible explanation of this phenomenon is, according to Kimura, that dichotic presentation demands more from the assimilation mechanism resulting in a more marked effect.

In the frame work of a research programme we were able to test the ear asymmetry hypothesis and to make the age trend of this variable the subject of research.

Method

Subjects

One hundred and twenty children took part in the project: 60 girls individually matched on the basis of age with 60 boys. There were 20 children of 6 years old, 20 of 7 years old, etc., up to 20 of 11 years old. The children came from normal primary schools.

Tests and apparatus

A total of 18 series of digits and 18 sound patterns were presented to each ear separately. The right and left ear series were randomly selected. The digit series comprised four, five or six digits. The sound patterns were Morse-like series of dots and dashes, generated by a buzzer. These series comprised three, four or five elements. Digit series and sound patterns were recorded on a one channel tape recorder and could be relayed to each ear separately via an earphone. The reproduction of digit series was made orally and the reproduction of sound patterns was made with the aid of the buzzer.

Procedure

The sequence in which the children were tested was random. Alternately, one child was tested, using first the right ear, then the left ear, followed by a child who was tested using first the left then the right ear. The order of presentation of digit series and sound patterns was random, except that the shortest series (four digits) and smallest patterns

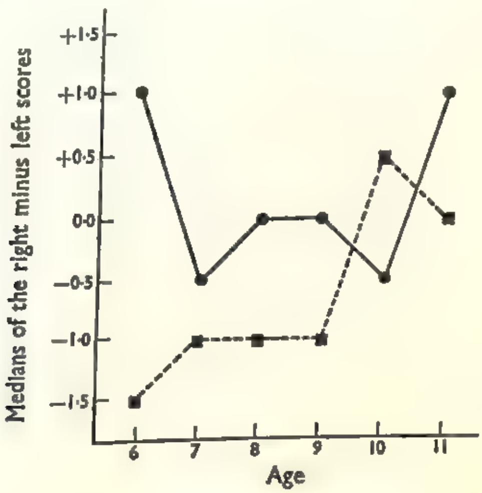
(three elements) were always presented first. The tester made certain that the testee understood the instruction by means of three training trials. A response was only scored as correct when a digit series or sound pattern was reproduced entirely correctly.

Results

The retention of digit series was not significantly better via the right than via the left ear ($z = 0{\cdot}51$, $p = 0{\cdot}30$ one-tailed). The retention of sound patterns was significantly better in the case of input via the left ear than via the right ear ($z = 2{\cdot}01$, $p = 0{\cdot}02$ one-tailed). Fifty-three children (28 girls and 25 boys) obtained better results in reproducing digit series presented via the right ear, 47 children (22 girls and 25 boys) obtained better results via the left ear. This difference is not significant ($z = 0{\cdot}50$, $p = 0{\cdot}31$ one-tailed). With the reproduction of sound patterns 59 children (32 girls and 27 boys) obtained better results via the left ear, 38 children (17 girls and 21 boys) better results via the right ear. This difference in numbers of children, who achieve better results in the reproduction of non-verbal material presented via the left than via the right ear, is significant ($z = 2{\cdot}04$, $p = 0{\cdot}02$ one-tailed).

As can be seen from the above mentioned figures, a marked difference between the results of boys and girls was not found.

FIGURE 1



- - - non verbal,——verbal

In Figure 1 age trends are shown for the medians of the intra-individual right minus left differences. A nonparametric trend analysis (Ferguson, 1965, Ch. 10) gave for verbal material a very significant quadratic component ($z = 3{\cdot}42$, $p = 0{\cdot}0007$ two-tailed). For non-verbal material a linear component was found, approaching the 5 per cent. level of significance ($z = 1{\cdot}95$, $p = 0{\cdot}051$ two-tailed).

Discussion

Children of primary school age perceive and retain non-verbal material markedly better by means of the left ear than via the right ear. This result was obtained in spite of the fact that the stimulation of the ears took place separately, at separate moments. Moreover, this finding is in accordance with Milner's (1962) view that the right temporal lobe is dominant in the auditory perception of non-verbal material. Contrary to expectations the verbal material produced no better results when presented to the right ear than to the left ear. It is conceivable that the verbal material (digit series) was too simple compared with the non-verbal material (sound patterns) so that too small an appeal was made to the assimilation mechanisms responsible for ear asymmetry. In this context we refer to a recently published research of Shankweiler and Studdert-Kennedy (1967), which clearly shows that results are dependent on the nature of the material.

It is of special interest to note the relationship to age which we found in this investigation. With non-verbal material better results are achieved at first age level by the left

ear. With increasing age however this changes in favour of the right ear. Using verbal material better achievements are obtained at an earlier age with the right than with the left ear. This changes fairly soon in favour of the left ear until after the age of 10 when the right ear takes the lead again. On the basis of this investigation it cannot be established whether the separating trends after the age of 10 are continued later in age or not. Should this be so, then the age of 10 could be regarded as the critical stage at which ear asymmetry becomes consistent. Such critical ages have been found for other aspects of lateral dominance (Belmont and Birch, 1963).

We should like to thank colleagues of the Paedological Institute Research Department (address: Prins Hendriklaan 23, Amsterdam), who have contributed to the establishment of this research project.

References

Belmont, L., and Birch, H. G. (1963). Lateral dominance and right–left awareness in normal children. *Child Developm.*, **34**, 257–70.

Broadbent, D. E., and Gregory, M. (1964). Accuracy of recognition for speech presented to the right and left ears. *Quart. J. exp. Psychol.*, **16**, 359–60.

Bryden, M. P. (1963). Ear preference in auditory perception. *J. exp. Psychol.*, **65**, 103–5.

Ferguson, G. A. (1965). *Nonparametric Trend Analysis.* Montreal: McGill University Press.

Gesell, A., and Ames, L. B. (1947). The development of handedness. *J. genet. Psychol.*, **70**, 155–75.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canad. J. Psychol.*, **15**, 166–71.

Kimura, D. (1964). Left–right differences in the perception of melodies. *Quart. J. exp. Psychol.*, **16**, 355–8.

Milner, B. (1962). Laterality effects in audition. In Mountcastle, V. B. (Ed.), *Interhemispheric Relations and Cerebral Dominance.* Baltimore: Johns Hopkins Univ. Press.

Palmer, R. D. (1964). Development of a differentiated handedness. *Psychol. Bull.*, **62**, 257–72.

Shankweiler, D., and Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to left and right ears. *Quart. J. exp. Psychol.*, **19**, 59–63.

Manuscript received 2nd June, 1967.

KINAESTHETIC JUDGEMENTS OF THE DIRECTION OF LINE BY YOUNG CHILDREN

BY

RAY OVER and JOAN OVER

From the Psychology Department, University of Otago

The ability of young children to make kinaesthetic judgements of direction of line has been studied under recognition and detection conditions. Some children were not more accurate than chance when they were required to indicate the same one of a pair of mirror image oblique lines over a series of trials; these children could, however, detect whether two obliques were oriented in the same direction or in different directions. These latter data suggest that the poor recognition of mirror image obliques arises from deficiencies in spatial categorizing and not in input coding. This conclusion is supported by the finding that children who performed poorly on the kinaesthetic recognition of obliques performed poorly on the visual recognition of these figures.

Introduction

The visual discrimination of shape by young children has generally been studied by presenting a child with two shapes and requiring the child, on each of a series of trials, to point to the shape designated by the experimenter as "correct." It has been found that the young child (Rudel and Teuber, 1963; Over and Over, 1967), like the octopus (Sutherland, 1957), does not perform this discrimination at a better than chance level when the two shapes are mirror image oblique lines. Both species, however, readily discriminate between a vertical and a horizontal line.

Data on the discriminability of shapes have been used to develop inferences about the operating characteristics of neural mechanisms by which input from a shape is analysed (Dodwell, 1957; Deutsch, 1962; Sutherland, 1963). It has been suggested (Over and Over, 1967), however, that it may not be valid to interpret data obtained by the method described above in this way in that poor discrimination may occur through the subject being unable to remember from trial to trial the spatial properties which define the shape associated with reinforcement rather than through the shapes generating similar signals in the visual system. With visual measures it has been found (Over and Over, 1967) that many young children who are unable to discriminate between mirror image obliques when tested by the method described above (hereafter called the *recognition* method) are readily available to *detect* whether two obliques line are identical or different in orientation. Accordingly, poor discrimination under recognition conditions cannot be interpreted as indicating that mirror image obliques looked to be similarly oriented to the children.

If the poor visual recognition of mirror image obliques by young children arises from limitations in the system by which information relating to spatial orientation is remembered, it might be expected that young children would find the kinaesthetic recognition of mirror image oblique lines more difficult than the kinaesthetic recognition of the vertical and horizontal. They should, however, readily be able to judge whether two obliques inspected kinaesthetically are identical or different in their orientation. These expectations are examined in the present experiment.

Method

Subjects

Sixteen children, whose ages ranged from 5 years 6 months to 6 years 5 months (mean 6 years 1 month), were used as subjects. The children were selected from a public elementary school and were rated by their teacher as being of average ability. There were eight boys and eight girls.

Apparatus

Each shape used to obtain visual judgements was made by attaching a black plastic strip, 3 in. long and ¼ in. wide, to a 5-in. square of white opaque Plexiglas. Wooden strips, 3 in. long, ¼ in. wide and ¼ in. high, and glued to hardboard squares, were used as stimuli for kinaesthetic discrimination.

Procedure

Each child was initially required to make kinaesthetic judgements. Half of the children were tested under recognition conditions before being tested under detection conditions, and within each method half of the children were tested with the horizontal and vertical before being tested with the mirror image obliques. Each stage of testing continued until the subject either was correct in his judgements on nine successive trials or had completed 40 trials without reaching this criterion.

The children were not blindfolded but were required to place both hands beneath a curtain and to feel the shapes. Under both testing methods two shapes were always present and the child was allowed to move one hand over one shape and the other over the other shape before making a judgement. The centres of the two shapes were always 8 in. apart. Under the recognition conditions the two shapes always differed in their orientation (one was horizontal and the other vertical, or they were mirror image obliques). The child was told that selecting one of the shapes would always be "right" and selecting the other "wrong" and that his task was to find the correct shape and to indicate it on each trial. After each judgement the child was told whether his choice was correct. The relative positions of the two shapes were varied over trials such that a given position was not correct for more than two successive trials. For each pair of shapes, one figure was designated "right" for half of the subjects, and the other figure was "right" for the remaining subjects.

Under detection conditions the two shapes were of the same orientation (e.g. both vertical) on half of the trials and were differently oriented (e.g. one horizontal and the other vertical) on half of the trials. The child was required to judge whether the two shapes were the same or different and on each trial he was told whether his judgement was correct. The sequence of presentation of the stimuli was arranged so that a given response category was not correct for more than two consecutive trials.

After he had completed the kinaesthetic judgements each child was tested for the visual recognition of mirror image obliques. Two plaques displaying the shapes were presented 2 in. apart and their relative positions over trials were varied in the manner described above. One of the obliques was designated "right" for half of the children and the other oblique was correct for the remainder. On each trial the child was told whether his judgement was correct. Forty trials were given unless the criterion of nine successive correct trials had earlier been reached.

TABLE I

Recognition and Detection Measures of Discrimination

		Horizontal and vertical		*Mirror image obliques*	
		Mean	*SD*	*Mean*	*SD*
Kinaesthetic recognition	F	1		7	
	P	12·33	5·94	16·11	7·01
	E	2·69	5·00	9·94	8·09
Kinaesthetic detection	F	0		1	
	P	13·38	5·48	18·87	8·41
	E	1·94	2·51	5·00	4·72
Visual recognition	F			6	
	P			12·40	3·98
	E			6·56	7·05

F—subjects failing to reach criterion within 40 trials;
P—trials required by subjects who reached criterion;
E—errors made by all subjects.

Results

Table I shows, for the two sets of shapes and the two methods of judgement, the number of children who failed to reach the criterion, the mean number of trials required by subjects who attained the criterion, and the mean number of errors made on each trial by all subjects. Data on the visual recognition of mirror image oblique lines are also shown.

Only one child was unsuccessful on the kinaesthetic recognition of the horizontal and the vertical, and no child was unable to judge whether a horizontal and a vertical line were similarly oriented. Seven of the 16 children failed to reach the criterion of discrimination on the kinaesthetic recognition of mirror image oblique lines; only one of these five children was, however, unable to judge whether two oblique lines were oriented in the same or in different directions. These differences between recognition and detection data are similar to those found for visual discrimination (Over and Over, 1967). A close relationship was found between measures of the kinaesthetic recognition and the visual recognition of mirror image obliques. All subjects who were successful on kinaesthetic recognition were also successful on visual recognition. Six of the seven subjects who were unsuccessful on kinaesthetic recognition were also unsuccessful on visual recognition. Subjects who failed to reach the criterion of discrimination for a set of figures were generally correct on only half (mean 22·93, σ 2·58) of the 40 trials. Only one subject was correct on more than six successive trials on a task on which he failed to reach the criterion.

Similar trends are seen when an analysis, by t tests for correlated means, is made of the "errors" data shown in Table I. Under recognition conditions significantly more errors ($t = 3{\cdot}21$, $d.f. = 15$, $p < 0{\cdot}05$) were made in judging the oblique lines than the horizontal and vertical lines. Although Table I indicates that subjects who reached the detection criterion with both sets of shapes made more errors with oblique lines, the difference between the two sets of shapes in mean errors made is not significant ($t = 1{\cdot}57$, $d.f. = 14$, $p > 0{\cdot}05$). Significantly fewer errors ($t = 2{\cdot}18$, $d.f. = 15$, $p < 0{\cdot}05$) were made under detection than recognition conditions with the oblique lines. For the horizontal and vertical lines the mean difference in errors between recognition and detection conditions was not significant ($t = 0{\cdot}80$, $d.f. = 15$, $p > 0{\cdot}05$).

Significantly fewer errors ($t = 2{\cdot}62$, $d.f. = 15$, $p < 0{\cdot}05$) were made on the visual than the kinaesthetic recognition of mirror image obliques, and fewer trials were required to reach the criterion. It should be noted, however, that all subjects made visual judgements after completing kinaesthetic judgements. The difference in results may reflect the effects of prior practice and may not indicate that children find these shapes more difficult to discriminate in the kinaesthetic than in the visual modality. The present experiment has been primarily concerned with the kinaesthetic discriminability of different shapes and does not provide data on the above issue.

Discussion

Young children find the kinaesthetic discrimination of mirror image oblique lines to be more difficult than the kinaesthetic discrimination of the vertical and horizontal under recognition, but not under detection, conditions. While seven of the 16 subjects did not exceed chance accuracy on the recognition of obliques, only one failed to reach the criterion of nine correct judgements in succession within 40 trials when tested on detection with these shapes. Poor performance under recognition conditions thus cannot be interpreted as indicating that mirror image oblique lines felt to be similarly oriented to the subjects.

Successful performance under recognition conditions has accordingly involved something more than being able to feel that the shapes presented differed in orientation; the ability to remember the defining characteristics (the spatial orientation) of the "correct" shape from trial to trial also was necessary. The finding that subjects who perform poorly on kinaesthetic recognition with mirror image obliques also perform poorly on visual recognition with these shapes indicates that the discrimination of direction of line under recognition conditions is controlled by processes which are modality-general rather than modality-specific. It is possible that young children categorize and remember directional information by using the body as a frame of reference. Many young children find it difficult to remember which side of the body is left and which side is right but experience no such difficulty in remembering up (the head) and down (the feet).

Rudel and Teuber (1963) have suggested that the same processes may underlie the poor visual recognition of mirror image oblique lines by the young child and the octopus. The inability of the octopus to discriminate between obliques (Sutherland, 1957) has generally

been attributed to the operating characteristics of input coding and analysing mechanisms in the visual system and brain of the octopus (Dodwell, 1957; Deutsch, 1962; Sutherland, 1963). Data on the visual detection of obliques (Over and Over, 1967) and the present finding that there is a close relationship between visual and kinaesthetic judgements of direction indicate that the poor visual recognition of obliques by young children cannot be explained in these terms. It is possible that the poor visual recognition of mirror image obliques by the octopus also arises from deficiencies in the system by which spatial information is remembered. Whether this is so can be determined by comparing recognition and detection measures of discrimination. Unfortunately all available data on the discrimination of shape by the octopus have been obtained by use of recognition procedures.

References

Deutsch, J. A. (1962). A system for shape recognition. *Psychol. Rev.*, **69**, 492–500.

Dodwell, P. C. (1957). Shape recognition in rats. *Brit. J. Psychol.*, **48**, 221–9.

Over, R., and Over, J. (1967). Detection and recognition of mirror image obliques by young children. *J. comp. physiol. Psychol.* (in press).

Rudel, R. G., and Teuber, H.-L. (1963). Discrimination of direction of line in children. *J. comp. physiol. Psychol.*, **56**, 892–8.

Sutherland, N. S. (1957). Visual discrimination of orientation by octopus. *Brit. J. Psychol.*, **48**, 55–71.

Sutherland, N. S. (1963). Shape discrimination and receptive fields. *Nature* **197**, 118–22.

Manuscript received 6th June, 1967.

PERCEIVED DISTANCE IN IMAGINED SPACE

BY

WILLIAM EPSTEIN

From the Department of Psychology, University of Kansas

An experiment was designed to determine whether, in the absence of distance-cues, the perceived distance of an object is regulated by the imagined space in which the object is located. Judgements of distance and size were obtained in a completely reduced situation. Preceding these judgements, the subject engaged in 5 min. of visual exploration of the interior of an oblong box. The length of the box differed (3, 6 or 9 ft.) for different groups of subjects. The different pre-exposure conditions were expected to produce differences in the dimensions of the imagined space that the subject introduced in the subsequent test session. These differences in imagined space should lead to differences in the perceived distance of a single object located at a constant distance. Significant differences were obtained in the expected direction. In addition, judgements of the size of the standard tended to co-vary positively with the distance-judgements, suggesting that the effect of imagined space on judged distance was not simply the product of experimental biasing of the response system. Some methodological and theoretical implications of the findings are considered.

Introduction

At what distance does a single object appear to be when all of the recognized determinants of perceived distance have been eliminated? For example, when a nonrepresentative fluorescent figure is presented in total darkness, and is viewed monocularly from a distance that exceeds the effective distance for the accommodation reflex, the figure will appear to be localized somewhere in space. It will not be mistaken for an after-image, nor will it appear to wander aimlessly in space, changing distance at each moment of observation. Instead, a subject who is questioned in this situation will report that he sees a figure at some specific distance from himself. This paper offers an explanation of this observation, and reports an experimental test of the explanation.

The following hypothesis is proposed: In a totally dark room, a subject will introduce an imagined space, and localize all objects at an average distance determined by the dimensions of the imagined space. If, as in our example, a figure is exposed in a completely dark room, the subject will imagine or assume the dimensions of the unseen room, and he will localize the figure at a distance that approximates the midpoint of the room.

One procedure that can be used to examine this hypothesis is to vary the dimensions of imagined space, and observe the effects of the perceived distance of a standard that is located at a constant physical distance. In order to implement this procedure, an assumption was made about the variables that control imagined space. We assumed that the dimensions of imagined space are influenced significantly by the dimensions of the space that is directly viewed prior to the test. (The studies of the moon illusion by Gruber, King and Link (1963) support this assumption.) Therefore, an experiment was designed to manipulate imagined space by manipulating the dimensions of the immediately preceding visible space. It was expected that perceived distance and perceived size would vary systematically as a function of the different pre-exposed spaces.

Method

Subjects. Sixty undergraduate students were assigned to three experimental conditions in a regular rotation based on their arrival in the laboratory.

Apparatus. An oblong box was constructed, 10 ft. long and 20 in. in width and height. The front end of the box was open, and the location of the back wall could be varied at will. The floor of the box was covered with a regular, checkerboard, black and white, except for vertical stripes along the left-hand wall. The stripes demarcated 1-ft. intervals. Twenty-five numbered circular markers were distributed about the box, attached to the interior of the walls and the floor. The numbers were randomly selected from the numbers 1 to 99, and they were randomly assigned to locations in the box.

A chin-rest and an arrangement for immobilizing the head were located at the front end of the viewing box. When the subject's head was in the chin-rest, he could view the

interior of the box through red filters. Vision was monocular, due to a shield that blocked the view of the subject's left eye. Two alternative lighting conditions were provided: (1) normal incandescent illumination, and (2) ultraviolet illumination. In the latter case, only fluorescent surfaces could be seen through the red filter.

In addition to the viewing box, a set of size-comparison stimuli were prepared. Seven aluminium circular discs were glued to a smooth oblong strip of wood. The discs ranged in size from 13 to 37 mm. in steps of 4 mm. They were arranged in a row, in order of size. Adjacent discs were separated by interspaces of 1 in.

Procedure. The subject entered the room with his eyes closed. He was seated at the front of the box, and remained in total darkness for 2 min. Then, the incandescent light was turned on, and he looked into the interior of the well-illuminated box. For the next 5 min. the subject performed a target-location task. The experimenter called out the number of a marker, and the subject located the marker and described its location. In order to maintain the subject's alertness and guarantee full exploration of the box, the experimenter included several numbers that were not actually present in the box. The task was the same for all three experimental groups. Only differences between the lengths of the viewing-box distinguished the groups. By adjusting the position of the back wall, three lengths were obtained: 3, 6 and 9 ft.

Following the 5-min. exposure period, the incandescent lights were turned off, and the ultraviolet tube was turned on. In addition the subject's head was immobilized. Under these conditions, he made two kinds of judgements. He gave a verbal estimate of the distance from his eye to a fluorescent disc. The disc was 25 mm. in diameter, and was located at a distance of 60 in. at eye-level. Following the distance-judgement, the subject judged the size of the fluorescent disc. In reporting the size-judgement, the subject used the series of comparison discs. The experimenter described the arrangement of the discs, and informed the subject that the second smallest was the size of a dime, while the second largest was the size of a half-dollar. However, the subject was clearly informed that the standard disc was *not* intended to represent a coin. He was also explicitly permitted to reject all the choices and offer his own, e.g. "smaller than the smallest comparison disc." He was instructed to make a tactual match of his visual size-impression. The comparison-series was never available for visual inspection. The match was selected entirely by touch.

The instructions solicited reports of the subject's phenomenal impressions of distance and size. He was asked to report "the distance at which the disc looked to be in immediate perception. Base your response on your first impression of distance." Analogous phrasing was used in requesting the size-judgements.

Results

Distance-judgements. The mean distance judgement for the 3, 6 and 9 ft. pre-exposure conditions was 25·40 in. (SD = 11·23), 42·90 in. (SD = 13·71) and 72·45 (SD = 19·37), respectively. The difference between the means is highly significant (F, $2/57 = 34{\cdot}06$, $p < 0{\cdot}001$). Of the 20 subjects who were pre-exposed to the 36-in. box only one estimated the distance of the standard to be greater than 36 in. Of the 20 subjects who were pre-exposed to the 72 in. box, none estimated the distance to exceed 72 in., however there were 12 estimates that exceeded 36 in. Of the 20 subjects pre-exposed to the 108 in. box, only two subjects gave distance-estimates exceeding 108 in., however, there were a total of nine estimates that exceeded 72 in.

Size-judgements. The mean size of the comparison chosen to match the standard was 17 mm. (SD = 4·12), 21·80 mm. (SD = 5·83) and 25 mm. (SD = 6·63) for the 3 ft., 6 ft. and 9 ft. pre-exposure conditions, respectively. The difference between the means is significant (F, $2/57 = 10{\cdot}51$, $p < 0{\cdot}01$).

Relationships between size and distance. If the experimental treatments produce genuine modifications of perceived distance, then systematic variations of perceived size should accompany the differences in the distance-estimates. Although the data for individual subjects were often inconsistent, the general trend of the treatment means was for increases in judged size to be associated with increases in judged distance.

Discussion

The chief objective of this experiment was to show that in the absence of the recognized determinants of perceived distance, the localization of an object may be significantly influenced by the dimensions of imagined space. Imagined space was manipulated by

varying the length of the pre-exposure alley, with the expectation that distance-judgements would exhibit corresponding differences. The results confirmed this expectation. The fact that the variations in judged distance were accompanied by systematic variations in judged size, increases the likelihood that the distance-effects are localized in the perceptual system, and are not simply artefacts of induced response bias. Admittedly this latter argument is not conclusive. However, even if we accept the more conservative conclusion that the dimensions of imagined space merely serve to establish response preferences, there would be certain implications of importance.

The most obvious implication is for the methodology of experiments that arrange dark-room conditions to eliminate the visual cues for distance. The validity of this procedure will depend on the care that has been exercised to nullify or assess the effects of imagined space. There has always been an unarticulated recognition of this requirement, which is reflected in the widespread practice of preventing visual exploration of the laboratory room prior to the experiment, e.g. a subject is blindfolded before entering the room. But this procedure may only lead to uncontrolled effects of imagined space, rather than the nullification of these effects. In many cases, a more desirable procedure may be to expose all subjects to a common pre-test space, thus bringing the dimensions of imagined space under experimental control. This control will allow the effects of the experimental variable, e.g. variations of known size, to be exhibited more clearly. Gogel's (1964) review of the investigation of size as a cue to distance provides several examples of studies that might be improved by taking into account the effects of imaginary space.

The findings may also have relevance for a question that arises in evaluating the size-distance invariance hypothesis (Epstein, Park and Casey, 1961). What accounts for the visual angle matches that are obtained when all of the conventionally designated distance-cues have been eliminated? A number of investigators (e.g. Gogel, 1964, 1965; Wallach and McKenna, 1960) have contended that the retinal image size of a single object conveys no information about its size, unless some determinant of perceived distance is operative. Therefore, in explaining the classical Holway-Boring results for the completely reduced condition, Gogel and Wallach invoke an "equidistance tendency" or a subjective assumption of equidistance, which results in the apparent equidistance of the standard and the comparison. If a tendency towards equidistance prevails, then the visual angle matches can be derived from the size distance invariance hypothesis.* The present experiment illustrates one manner in which this tendency towards equidistance could come about. In experiments that involve a match between a standard exposed under reduced conditions and a comparison exposed under full-cue conditions, there will be a tendency for the unseen space surrounding the standard to assume the dimensions of the visible comparison space. The standard is then judged to be at a distance that is appropriate for the dimensions of the imagined space, thus insuring that the judged distance of the standard and the comparison will not differ greatly.

* Rock and McDermott (1964) challenge this analysis and offer an alternative.

This research was supported by Grant MH 4153 from the National Institute of Mental Health of the United States Public Health Service.

References

Epstein, W., Park, J., and Casey, A. (1961). The current status of the size-distance hypotheses. *Psychol. Bull.*, **58**, 491–514.

Gogel, W. C. (1964). Size cue to visually perceived distance. *Psychol. Bull.*, **62**, 217–35.

Gogel, W. C. (1965). The equidistance tendency and its consequences. *Psychol. Bull.*, **63**, 153–63.

Gruber, H. E., King, W., and Link, S. (1965). Moon illusion: an event in imaginary space. *Science*, **139**, 750–2.

Rock, I., and McDermott, W. (1964). The perception of visual angle. *Acta Psychol.*, **22**, 119–34.

Wallach, H., and McKenna, V. V. (1960). On size-perception in the absence of cues for distance. *Amer. J. Psychol.*, **73**, 458–60.

Manuscript received 15th May, 1967.

TRAINING OF FAST TAPPING WITH REDUCTION OF KINAESTHETIC, TACTILE, VISUAL AND AUDITORY SENSATIONS

BY

JUDITH I. LASZLO

From Department of Psychology The University of Western Australia

Five subjects were trained to tap on a light Morse-key during nerve compression block. The training sessions lasted for 40 sec., with a 5 sec. rest after the first 20 sec. work period. The group learning curve reached 89·5 per cent. level of normal performance by the eighth training session. In the ninth, the testing session, subjects tapped with visual and auditory sense reduction superimposed on the kinaesthetic and tactile impairment of the training condition. Performance in the testing session reached 40·9 per cent. of normal.

The sixth subject was trained in the same task as the other five subjects, but the training condition included elimination of cues from all four sensory channels. He reached 79·09 per cent. of his normal tapping performance in the seventh session.

These results show that the motor skill of tapping can be relearned in the absence of kinaesthetic cues. Furthermore when the subject has no conscious knowledge of any peripheral sensory cues connected with the ongoing motor activity, learning can nevertheless take place. These findings lead to the hypothesis, that skilled motor activity can be monitored by central processes alone.

During the training sessions subjects showed a tendency of tapping in groups of gradually increasing length. It is hypothesized that increased number of taps forming a group gives an indication to the possible mode of action of these central processes.

Introduction

In previous studies (Laszlo, 1966, 1967) it was shown that the loss of kinaesthetic sensation in the lower arm and hand caused a significant performance decrement in fast tapping on a Morse key. Learning to overcome the sensory loss was prevented by giving each subject one trial in any one condition only.

Taub, Bacon and Berman (1965) showed in a study on monkeys that the animal could learn an avoidance response in the absence of feedback from the forelimb. They argued that in rhizotomized animals, learning and performance of movement can, in fact, be mediated through central processes in the total absence of peripheral feedback from the active limb. Welford (1959) summarizing the relevant literature proposed that the effector mechanisms are capable of carrying out series of actions, treating the series as a single unit. These units may be composed of simple movements such as pressing or releasing a Morse key, or even triggering off a number of taps as a single unit. The effector mechanisms involved have both central and peripheral components. Further, the action of the effector mechanisms can be independent of input stimuli generated by the response.

The following study was designed to test these hypotheses directly. Thus the effect of training, under conditions of sensory deprivation, in tapping skill was investigated. Specifically, the following questions were raised:

(*a*) Can a skilled motor act be relearned when a subject is trained under the condition of kinaesthetic sense reduction?

(*b*) If training was effective, to what extent can a subject tap with exteroceptive loss superimposed on kinaesthetic and tactile sensory reduction?

(*c*) Can relearning of tapping take place if a subject is trained with sense impairment of all relevant modalities?

Method

Nerve compression block was applied to the arms of the preferred hand in six subjects to reduce kinaesthetic sensation. Five of them were trained to tap as fast as possible

on a light Morse-key, with visual and auditory feedback intact for eight training sessions. In the ninth session, visual and auditory sense reduction was added to kinaesthetic impairment. The sixth subject was trained in the same task but, in his case, throughout all training sessions sensory cues from kinaesthetic, tactile, visual and auditory channels were eliminated. Tactile sensation was lost in all subjects in the course of the compression block.

The weekly training sessions lasted for 40 sec., with 5 sec. rest after the first 20 sec. work period. Performance was recorded on a modified Electrocardiogram.

Subjects. Three male and three female subjects were used. Two were members of the staff and four were students of the University of Western Australia.

Apparatus. Sphygomomanometer, light Morse-key, modified Electrocardiogram, recorder, earphones, masking tone generator and stopwatch were used.

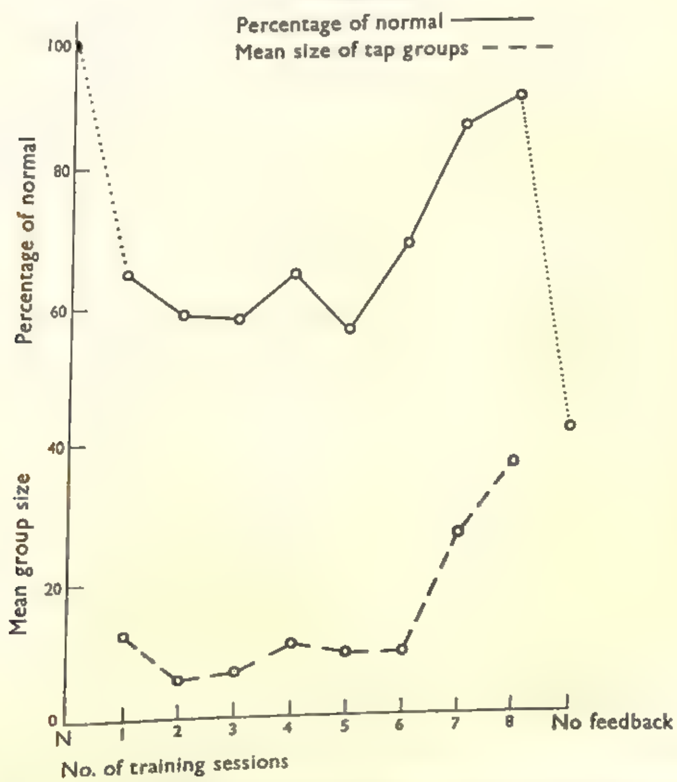
Procedure

Subjects were tested in individual sessions. In the preliminary session each subject was asked to tap on the Morse-key for 20 sec. as fast and as evenly as possible to obtain his normal record. Compression block was then applied, to familiarize the subject with the procedure. The compression block procedure is described in detail in a previous paper (Laszlo, 1966).

Each subject came once a week at a set time. The first five had eight training sessions and an additional testing session. In each of the training sessions, the subject underwent compression block procedure, and when kinaesthetic feedback was no longer perceived, tapped for 20 sec. on the Morse-key. Without shifting his finger off the key, he rested for 5 sec., and tapped for a further 20 sec.

In the testing session procedure was similar to the training sessions, except in regard to sensory control. Here, in addition to the kinaesthetic and tactile sensation loss, visual cues were eliminated by shielding the subject's hand from his line of vision, and auditory feedback was cut out by a masking tone, applied through a pair of earphones.

FIGURE 1



Performance of five subjects trained without kinaesthetic sensation.

The sixth subject had seven training sessions. These were identical in procedure to the *testing* session of the other five subjects.

The score for each session was taken as the number of taps during a 20 sec. work period. Either the first or the second work period score was recorded, depending on which of the two gave a higher score. The higher score was chosen, as this gave a closer indication of the learning achieved in any particular session. The small variation between the first and second 20 sec. period could possibly be caused by motivational fluctuations.

Additional information was gained by counting the taps that followed each other closely, i.e. forming a group of taps. The mean group size was then calculated.

Results

The results for the first five subjects are expressed as mean performance decrement in Figure 1.

The learning curve for the training sessions shows that performance reached 89·5 per cent. of normal level by the eighth trial. Performance dropped to 40·9 per cent. of normal in the testing session.

The lower curve on Figure 1 shows the mean number of taps grouped in each training session. Its slope follows the learning curve with slight departures.

A trend analysis on the mean group size data was carried out to test whether the increased grouping of taps gives a significant trend as training progresses and performance improves. The calculated F value of 1·386 is not significant.

TABLE I

TREND ANALYSIS OF MEAN TAP GROUP SIZE OVER EIGHT TRAINING SESSIONS FOR FIVE SUBJECTS

Source of variance	*Sum of squares*	*d.f.*	*Mean square*	*F*
Trials (T) ..	3944·069	7	563·438	1·386
Subjects (S) ..	2571·216	4	642·804	
S × T	11385·897	28	406·639	
Total ..	17901·182	39		

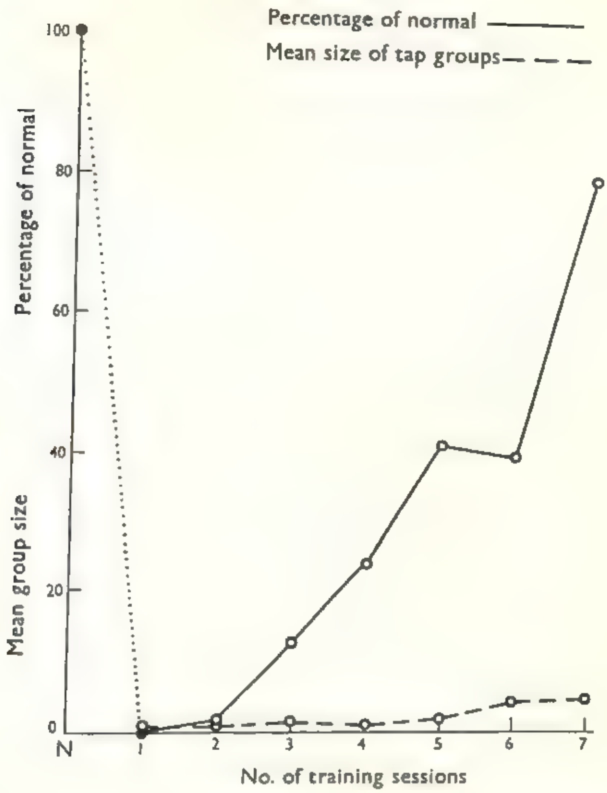
A second trend analysis, omitting the first trial was then performed. In view of the J shaped curve, it was argued that possibly a higher motivational level could have influenced the performance in the first trial, as compared with the rest of the training trials. The F value, 5·83 is significant on the 0·01 per cent. level of confidence (F value from table of F values = 3·67 at 0·01).

TABLE II

TREND ANALYSIS OF MEAN TAP GROUP SIZE OVER SEVEN TRAINING SESSIONS FOR FIVE SUBJECTS

Source of variance	*Sum of squares*	*d.f.*	*Mean square*	*F*
Trials (T) ..	8957·11	6	1492·85	5·83
Subjects (S) ..	2421·43	4	605·36	
S × T	6148·41	24	256·18	
Total ..	17526·95	34		

FIGURE 2



Performance of one subject trained without peripheral feedback.

The results for subject 6, who was trained in fast tapping in the absence of kinaesthetic, tactile, visual and auditory sensations are presented in Figure 2. This shows that training in the absence of all perceived sensation did give a performance of 79·09 per cent. of normal. The mean group size curve does not follow the learning curve as closely as in Figure 1, but still shows a trend of increase in group size.

Discussion

The results show that in the absence of kinaesthetic sensation subjects undertaking eight training periods can reach 89·50 per cent. of tap numbers recorded under normal conditions. One subject trained with peripheral sense impairment in four modalities reached 79·09 per cent. of normal fast tapping in the seventh training session. It was also found that with the advancement of training, taps were grouped into groups of gradually increasing size.

It was shown previously (Laszlo, 1967) that, in the absence of kinaesthetic sensation, cutting out a single exteroceptive channel does not result in further performance decrement in fast tapping. Loss of both visual and auditory sensation, with intact kinaesthetic sensation, does not impair performance at all. Thus the loss of kinaesthesis was found to cause the greatest performance decrement when compared with the other modalities tested. When all sensory channels were blocked there was further substantial impairment in performance. The question arises whether, in training to tap in the absence of kinaesthetic information, a subject learns to rely increasingly on exteroceptive modality or to recruit central processes. Inspection of Figure 1 shows that in the testing session where all sensory cues were eliminated, performance dropped to 40·9 per cent. of normal. This decrement is considerably less than the decrement shown by the 10 subjects in the previous study (Laszlo, 1967)—i.e. to 29·0 per cent. of normal. These 10 subjects were not trained. The difference from 29·0 to 40·9 per cent. performance in the absence of

relevant feedback for the two groups supports the notion that central processes are involved in the monitoring of the skill under this condition.

Looking at the learning curve in Figure 2, the increase in the number of taps as training proceeds cannot be explained by any other but central factors.

Only one subject was used here as the question raised—i.e. is relearning of a skilled motor task possible without peripheral feedback?—was answered by the results gained (Dukes, 1965).

The subject had no knowledge of his performance at any of the work periods. In fact, as soon as he stopped work he asked whether he tapped at all. He told the experimenter that he was giving himself instructions to "lift and push" his finger. When asked to say the instructions aloud in Session 4 it was observed that the instructions did not coincide with his tapping. These observations show that the subject was neither consciously perceiving sensation from the periphery, nor controlling movement during these training periods.

While no neurophysiological knowledge of the exact mechanism of the central processes is as yet available, the present study gives some possible indication as to the mode in which these processes enhance learning. The mean group size curve in Figure 1 shows a near parallel increase in the mean size of the tap groups and performance increment. As central processes seem, in fact, to be involved here, it is possible that the size of the "units" (Welford, 1959) of the skill are increased as learning progresses. That is, at the commencement of training under impoverished sensory conditions tapping is monitored as single taps or even in "lift" and "push" units. As training advances, however, groups of two taps are triggered off, then groups of three and so on. This possibility is supported by the evidence from Figure 2 also. Here mean group size is increased in the sixth and seventh sessions. Performance increment was noticeable from the third session, with decreased intervals between tap groups.

In Figure 1 both learning and mean group size curves approximate a J shape. The trend analysis supports the significance of increase in group size, but only from the second trial onward. It is not unlikely that the first trial falls outside the trend due to motivational factors.

It can be argued that Chase's information flow model (Chase, 1963) could not explain motor control in the absence of peripheral information. Improvement in performance under the conditions of the present study is even more difficult to account for within the framework of the model. The error detection unit cannot function without information received from the periphery. Without incoming information matching operations of ongoing activity to the standard is impossible. Consequently, error correction programming is also impaired and hence reduction of errors by effectors cannot be explained. Both error detection and error correction units are postulated as central nervous system mechanisms. To adapt the model to accommodate the present findings a further central unit would have to be incorporated, namely, an autonomous central control unit. The nature of the actual central mechanisms involved here is not known, but workers in the field of movement control do postulate central processes. Teuber (1964) talks of corollary discharge from the motor system to the sensory system prior to sensory feedback, and proposes that this discharge is in fact the basis of voluntary movement.

It is known that the sensory and motor areas of the central nervous system are closely interrelated. Rosenblith and Vidale (1962) state that our thinking on the separate role of sensory and motor structures has to be revised, in view of the extensive overlap in sensory and motor functions in the C.N.S. This interdependence of the two areas might well be the basis of central control of voluntary movement.

Another point of view regarding central mechanisms is expressed by Merton (1964). In his description of the "sense of effort," he claims that the organism is capable of "acquiring knowledge of events outside the central nervous system without the use of sense organs" (p. 398). Also that through the sense of effort the organism "knows" the extent of a movement in the absence of peripheral feedback. That the sense of effort is not consciously perceived is not a valid argument against its acceptance as "the code in which the central nervous system operates is never apparent subjectively" (p. 399). From Merton's viewpoint, it could be argued that in the absence of peripheral feedback, movement is "perceived" and controlled by the sense of effort, a central process.

This study gives empirical evidence to show that purely central processes can initiate and control a motor act. Further, the data also show that performance can improve with training when monitoring is dependent on these central processes alone. Evidence is not gained, however, to clarify how these central mechanisms become more efficient.

To say, as was found, that the "units" of the skill become increasingly larger can be taken as an indication of the mode of action of these possibly enhanced processes, but not as an explanation of the actual processes themselves.

References

Chase, R. A. (1963). Information system analysis of the organization of motor activity. In Hoch, P. H., and Zubin, J. (Eds.), *Disorders of Perception*. New York: Grune and Stratton.

Dukes, W. F. (1965). N = 1. *Psychol. Bull.*, **64**, 74–9.

Laszlo, J. I. (1966). The performance of a simple motor task with kinaesthetic sense loss. *Quart. J. exp. Psychol.*, **18**, 1–8.

Laszlo, J. I. (1967). Kinaesthetic and exteroceptive information in the performance of motor skills. *Physiology & Behavior*, in press.

Merton, P. A. (1964). Human position sense and sense of effort. *Homeostasis and feedback mechanisms. The 18th Symposium of Society for Experimental Biology.* London: Cambridge University Press.

Rosenblith, W. A., and Vidale, E. B. (1962). A quantitative view of neuroelectric events in relation to sensory communication. In Koch, S. (Ed.), *A Study of a Science*, Vol. 4, New York: McGraw-Hill.

Taub, E., Bacon, R. C., and Berman, A. J. (1965). Acquisition of a trace-conditioned avoidance response after deafferentation of the responding limb. *J. comp. physiol. Psychol.*, **59**, 275–9.

Teuber, H-L. (1964). Speech as a motor skill with special reference to non-aphasic disorders. In *Acquisition of Language Monographs of the Soc. for Research in Child Development*, **29**, 134–5.

Welford, A. T. (1959). Evidence of a single channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, **11**, 193–210.

Manuscript received 17th April, 1967.

THE PSYCHOLOGICAL REFRACTORY PERIOD AS A FUNCTION OF PERFORMANCE OF A FIRST RESPONSE

BY

MARILYN C. SMITH*

From the Department of Psychology, University of Toronto

Proponents of a "single channel" theory of the psychological refractory period have not specified whether the single channel occupies only the decision component of the response selection, only the motor or response component, or both. In this experiment, the delay in the RT to the second of two successively presented stimuli was examined as a function of whether or not an overt motor response was made to the first stimulus, keeping the decision component constant. It was found that in both conditions RT_2 was delayed, suggesting that the decision component was a part of the single channel. However, RT_2 was delayed by a significantly greater amount if a motor response *was* required, indicating that the motor component is part of the single channel as well. Implications of the results for an expectancy theory of the psychological refractory period are discussed.

INTRODUCTION

When two stimuli are presented in rapid succession, the reaction time (RT) to the second stimulus is usually delayed. This phenomenon is commonly referred to as the psychological refractory period, and various theories of the basis for the delay have been put forward (Smith, 1967). One explanation for the psychological refractory period which has recently received considerable support is that proposed by a "single channel" theory (Davis, 1956; Welford, 1959). According to this theory, somewhere in the processing mechanism there is a "channel" of limited capacity which can deal with only one of the two tasks at any given time. Consequently, the second response selection is delayed at this "bottle-neck," and processing must occur sequentially.

However, previous experiments have not specified just where in the system this "single channel" occurs. If response selection is considered to be composed of at least two parts—a decision component and a motor or response component—then theoretically the delay could occur at either or both these stages. It is the purpose of this experiment to determine which of these components are involved in the "bottle-neck." To do this, the decision component is kept constant, and delays in RT_2 are examined as a function of whether or not an overt response is performed to the first stimulus (S_1). If the bottle-neck occurs only at the motor component stage, then there should be *no* delay in RT_2 when no motor response is required to S_1, since this component would not be involved. On the other hand, if the single channel occupies only the decision component, then delays in RT_2 should be of the same magnitude regardless of whether a motor response is made, since the decision component is held constant in both conditions. Finally, if the single channel occupies *both* the decision *and* motor components, such that the second decision cannot begin until after the motor component of the first response selection is completed, then delays should be greater when a motor response must be made to S_1.

Experiments which have examined the delay in RT_2 as a function of whether or not the first response was performed have not provided a clear-cut answer. Davis (1959) found that while the delay in responding to the second of two stimuli was greater if a response was required to the first stimulus than if no response was required, most of the delay was still present even in the no response situation. However in a later study in which more stringent controls were employed (Davis, 1962), he found that if no response was required to the first stimulus, and if the inter-stimulus interval (ISI) was regular (i.e. the same ISI was employed during a block of 20 trials), then there was *no* delay at all in RT_2. Even with a random ISI, Davis found that if no response was required to the first stimulus there was a delay in RT_2 only at the shortest ISI of 50 millisec. Kay and Weiss (1961) have similarly found that RT_2 was not delayed as much when no response was required to the first stimulus.

However, in both these experiments no measures were taken to ensure that a decision was required to S_1 on those trials when no overt response was needed, or that decision

* Now on leave of absence at Department of Psychology, Carleton University, Ottawa, Canada.

times were the same regardless of whether the performance of a response was required to the first stimulus. When no overt response was required, subjects might simply have ignored S_1, or delayed processing the information until after the second stimulus had been responded to, as Davis himself suggests.

In order to avoid this problem, and to vary *only* whether or not a response had to be performed to S_1, a go no-go paradigm was employed, and delays were examined as a function of whether a first response had been made.

Method

Four male M.I.T. undergraduates served as subjects in this experiment. They focused upon a small focusing light, and were presented with two non-overlapping visual stimuli (illumination = 2·5 ft. candles) in rapid succession. The first stimulus was a luminous disc (visual angle = 0·68°) which could be presented either to the left or right of the focusing light. Subjects made a key-press response to this stimulus, pressing the left key if the disc was to the left of the focusing light and the right key if it appeared to the right. The second stimulus was a luminous bar (1·36°), which could appear either above or below the focusing light. A verbal response was made to this stimulus. Subjects were instructed to say "boo" if the bar was above the focusing light and "bun" if it was located below it. Both stimuli were presented for 20 millisec., and RTs for both the first and second responses were measured by two Hunter model 120A Klock-counters.

Five ISIs were employed: 25, 50, 100, 200 and 600 millisec. For two subjects (YK and EI), the ISIs were randomly presented. For the other two (DE and CH), the ISIs were constant over blocks of 20 (regular presentation).

Subjects always made the second verbal response (to the bar). However, the first manual response was made on only half of the trials. Subjects attended a total of 11 1-hr. sessions, of which the first constituted a practice session. For the first five experimental sessions the first stimulus (disc) was responded to only if it appeared to the left of the focusing light for two of the subjects and if it appeared to the right of the focusing light for the other two. In the last five sessions this was reversed. The left-right position of the disc stimulus was randomized. A total of 100 measures were taken at each ISI with the subject responding to the first stimulus, and 100 measures at each ISI with the first response omitted. To reduce competing responses, the subject removed the non-responding hand from the key and rested it in his lap. The data from the first session of each condition were not included in the analysis.

In addition, 24 two-choice verbal RT measures were taken each day, 12 measures at the beginning of the session and 12 at the end.

Results

The mean second RTs for each of the subjects under the two conditions of responding or not responding to the first stimulus are shown in Table I. The mean two-choice verbal RT is also shown to indicate the amount of increment when the disc was presented.

It is clear that under both the regular and random presentations of ISIs the second RT is greatly increased when the subject has to make a decision about the first stimulus (Is it left or is it right?) before he can attend to the second response, even if the performance of an overt response is not required. This delay averaged 195 millisec. at the shortest ISI.

A comparison of RTs under the conditions of responding or with-holding the first response, as shown in Table I, indicates that RT_2 was increased to a greater extent if subjects were required to perform a response to the first stimulus. A two-way analysis of variance, comparing response vs. no response conditions and decline in RTs over the range of ISIs employed, showed that both these variables were significant beyond the 0·005 level.

Discussion

This experiment was concerned with determining which components of a response selection occupied a "single channel" and therefore had to proceed sequentially, rather than in parallel. Our first finding, that RT_2 was delayed even when no overt response was required to S_1 indicates that the single channel must include the decision component of a response selection. In other words, the second response selection cannot begin until the decision component associated with the first response selection is completed. The second finding, that the delay is significantly greater under conditions which require the performance of an overt response to S_1, with the decision component held constant, suggests that the single channel extends to the motor component as well. Since the writing of this paper, similar results have been reported by Nickerson (1967).

The results of this experiment provide considerable difficulty for an "expectancy theory" of the psychological refractory period (Adams, 1962; Elithorn and Lawrence,

TABLE I

RT_2 AS A FUNCTION OF THE PRIOR PERFORMANCE OF RT_1

			Inter-stimulus interval (millisec.)					*Two-choice verbal RT*
			25	50	100	200	600	
Response required to S_1	Regular ISI	DE	699	721	668	616	533	430
		CH	566	576	531	479	380	375
	Random ISI	YK	1,042	1,024	1,038	963	776	680
		EI	612	622	598	547	516	425
		Mean	730	736	709	651	551	
No response required to S_1	Regular ISI	DE	638	636	616	629	553	
		CH	512	505	462	432	416	
	Random ISI	YK	932	946	967	897	818	
		EI	618	594	574	524	484	
		Mean	675	670	655	621	568	

1955; Kay and Weiss, 1961). This theory attributes the delayed second reaction not to a fundamental limitation in the system, but rather to subjective uncertainty about the time of occurrence of the second stimulus. This subjective uncertainty is totally determined by the characteristics of the ISI series employed, such as its range, frequency and duration. Since expectancy is dependent only upon the ISI series, and not upon the nature of the processing required for S_1, this theory would predict that delays in RT_2 should be the same under both the response and no response conditions, since the ISIs employed are identical. Our results indicate that this is clearly not so. If it is argued that the subject's expectancy for the second stimulus cannot begin until after the motor component of RT_1, rather than after the presentation of the first stimulus, then it is essentially being said that the subject acts as a sequential "single channel," processor, since preparation of RT_2 cannot be started until after completion of RT_1. In other words, the expectancy theory would have become a single channel theory.

This work was done at the Massachusetts Institute of Technology, and was supported by a grant to the author from that institution.

References

Adams, J. A. (1962). Test of the hypothesis of psychological refractory period. *J. exp. Psychol.*, **64**, 280–7.

Davis, R. (1956). The limits of the "psychological refractory period." *Quart. J. exp. Psychol.*, **8**, 24–38.

Davis, R. (1959). The role of "attention" in the psychological refractory period. *Quart. J. exp. Psychol.*, **11**, 211–20.

Davis, R. (1962). Choice RTs and the theory of intermittency in human performance. *Quart. J. exp. Psychol.*, **14**, 157–66.

Elithorn, A., and Lawrence, C. (1955). Central inhibition—some refractory observations. *Quart. J. exp. Psychol.*, **7**, 116–27.

Kay, H., and Weiss, A. D. (1961). Relationship between single and serial reaction time. *Nature*, **191**, 790–1.

Nickerson, R. S. (1967). Psychological refractory phase and the functional significance of signals. *J. exp. Psychol.*, **73**, 303–12.

Smith, Marilyn C. (1967). Theories of the psychological refractory period. *Psychol. Bull.* **67**, 202–13.

Welford, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, **11**, 193–210.

Manuscript received 22nd May, 1967.

LATENCY OF DIFFERENT VERBAL RESPONSES TO THE SAME STIMULUS

BY

PAUL FRAISSE

From the Laboratoire de Psychologie Expérimentale et Comparée de la Sorbonne, Paris

A response of reading (letter O) or of naming (circle) can be given to the same sign O. The verbal reaction time is higher when naming than when reading (difference 100 millisec.). This fact verifies that naming is a longer process than reading, the difficulty of perceiving the stimulus being equal.

The response "zero" which may be given to the same sign is nearer to the reading-response time than the naming-response. These facts can be explained if we say that uncertainty for coding concrete signs is greater than for alphabetical or numerical signs.

Introduction

In a series of experiments, Fraisse (1960) showed that a longer time was required to name geometric figures, colours, and drawings than to read the corresponding words, the uncertainty being kept equal. This fact, known since J. Mck Cattell (1886), is obtainable only in a choice reaction time situation, when one sign is chosen for naming or reading from among several others presented in a random order. On the other hand, it was found that time for naming increased proportionally with the uncertainty, when the time for reading was practically constant (Fraisse, 1964). In order to explain these results, we introduced the notion of compatibility. There would be less compatibility between a concret[illegible] and its verbal naming than between a word and its reading, which would mean that [illegible]ding when naming is more uncertain than coding when reading. In other words, when we are reading a word, the answer is relatively automatic, but when we are naming an object, there are different possible competing responses.

However, our investigations have revealed that the compatibility factor interacts with the discriminability of the stimulus, which is always an important variable in experiments on perception. When discussing our results with Professor Oldfield and his staff, it was suggested that we test our hypotheses in a particular case by using a graphic sign which might be encoded in different ways. In this case, differences in reaction time could not be attributed to the ease or difficulty of recognizing the stimulus. The letter O meets this requirement. It may be responded to in three ways: as a letter it may be *read*; or it may be seen as a *circle*, and then be named; finally, it may be seen as a *number*, namely, "zero." (For the subjects, who were French native speakers, to the symbol O in a numerical context, there is only one familiar answer which is "zero.")

If, when varying the modality of response (naming or reading) to the same stimulus, differences in reaction time are found, a further proof that naming and reading are dependent upon two different systems of coding will have been obtained.

Method

Subjects. Twelve students (six male and six female) were subjects.

Materials. A circle 15 mm. in diameter was drawn with indian ink. This figure was presented with three different series of stimuli, each of them being the same size as the figure or reference.

(*a*) four letters: O, B, F, K

(*b*) four numbers: 0, 2, 5, 7

(*c*) four geometric figures: circle, square, triangle, cross.

Procedure. An electronic Dodge Gerbrandt type tachistoscope was used to present subjects with the stimuli for durations exceeding that of reaction time, i.e. exceeding half a second.

Each series of stimuli was given separately and the subject was instructed as to which response to give. For instance, when presented with O in the series of letters, he knew he had to say "O"; in the series of numbers, to say "zero" and, in the series of geometric

figures to say "circle." The subject was always very familiar with the four alternatives in each series.

Each subject, in each series, was presented with the four stimuli 11 times in a random order. Hence, there are 11 responses of each sort: "O," "zero" and "circle." For statistical purposes the medians were used.

The order of presentation of the three series was randomized between the subjects.

In addition to the principal condition, each subject served under two control conditions. The first consisted of a measure of the choice reaction time in a reading situation. The letter "O", and the words "zero" and "circle" were presented successively to the subject in a random order. The letter E had been added to these three stimuli so that the uncertainty under this control condition should be the same as under the principal condition (four alternatives). This condition was introduced to ascertain whether the time for reading the three words differed significantly. Subjects gave 11 responses to this series of stimuli and the median for each word or letter was calculated.

The second condition was a simple reaction time. The three stimuli, the letter "O," and the words "circle" and "zero" were presented one after the other to the subject who was told which stimulus was to appear and who had to say as rapidly as possible the word or the letter. The aim of this condition was to determine the importance of the nature of the phonetic response on the recording apparatus in order eventually to take this variable into account when interpreting the results. Five successive reaction times were given by the subjects and again the medians were calculated.

Results

The means of the individual medians in the principal condition and in the control conditions are given in Table I.

TABLE I
Means of the Individual Medians (Millisec.)

		Responses		
		Letter O	*Zero*	*Circle*
Principal condition	Naming sign O	453	514	619
1st Control condition	Reading words in choice RT	495	507	511
2nd Control condition	Reading words in simple RT	304	334	333

An analysis of variance was performed. The results are shown in Table II.

In the principal condition, the difference between the latencies for response "circle" and the response "O" is very significant. This difference is 166 millisec. There is also a significant difference between the response "circle" and the response "zero": 105 millisec.

Thus, our main hypothesis is verified. The naming response takes longer than the reading response when reading the letter or the number.

Our control conditions confirm this result. Even if it is true that the answer "O" is about 30 millisec. briefer phonetically that the responses "zero" or "circle" (second control condition), this factor cannot explain the differences between the answer "circle" and the answer letter "O." Nor are the differences in the times for reading capable of explaining these differences since these times are very comparable (first control condition). In fact, the influence of the phonetic factor must affect these reading times and in view of this, the value 495 millisec. (RT for letter O) obtained in the first control condition would appear to be too high. This can be explained since we know that some subjects found it difficult to give the answer "O" as a response for reading. In the context of the first control condition, some subjects tend to give the response "zero" or "circle" instead of "O." However, it must be noted that in this control condition we drew *O* with the traditional

TABLE II

SUMMARY OF ANALYSIS OF VARIANCE

Source of variation	*d.f.*	*S.S.*	*M.S.*	*F*
Subjects	11	1926·54	175·13	4·34 < 0·01
Stimuli	2	1690·41	845·2	21·0 < 0·01
Figure-letter	1	1653·69	1653·69	41·1 < 0·001
Figure-number	1	654·38	654·38	16·3 < 0·01
Letter-number	1	227·55	227·55	5·65 < 0·05
Residual	22	885·78	40·26	
Total	35	4502·73		

ellipsoidal form. These wrong responses, when they were manifest, were eliminated, but it is possible that a slight inhibition of the response O has made the mean reaction time longer.

Finally, in our previous experiments with four alternatives, the difference in the responses for naming signs (or colours) was about 100 millisec. Results here are of the same order (108 millisec. for the response circle).

The response "zero" does not fit easily into our system of explanation. We consider it as a response of reading a sign analogous to the reading of a letter and not as a task of naming. However, it is slightly longer than the response "O." This difference can be partly explained on phonetic grounds, but also by the fact that response "zero" is less habitual and less frequent than response "O." It can be seen, however, that the response "circle" constitutes a particular category with a very different time of response, which the statistical analyses verify.

Conclusion

It can be concluded from this experiment that, given the same graphic sign to be read or named, the naming-response is longer than the reading-response (difference 100 millisec.), uncertainty being equal.

This result, obtained in the case where the graphic nature of the stimulus cannot play a role, agrees entirely with our previous results.

The longer time for naming cannot be explained either by the difficulty of perceiving the stimulus, or by phonetic characteristics of the response. It is at the level of the elaboration of the response, that is to say of coding, that the explanation is to be sought. We know that the time for naming increases with the uncertainty (number of alternatives) whereas the time for reading does not; so we are led to think that the essential factor in the extension of the time for naming comes from the uncertainty in coding or, in other words, from a smaller compatibility between a graphic sign and its naming than between a graphic sign and its reading. In order to explain this result, it is possible to put forward the fact that the relative discriminability of stimulus is not the same in each series. On the basis of our previous experiments, we may assert that the discriminability of the four stimuli is higher for geometric figures than for letters. All things being equal, the answers should have been shorter, but in this case, they were longer. We may thus conclude that naming an object is a more difficult coding operation than reading a linguistic sign.

References

CATTELL, J. MCK. (1886). The time it takes to see and name objects. *Mind*, **11**, 63–5.

FRAISSE, P. (1960). Recognition time measured by verbal reaction to figures and words. *Percept. mot. Skills*, **11**, 204.

FRAISSE, P. (1964). Le temps de réaction verbale: Dénomination et lecture. *Année Psychol.*, **64**, 21–6.

Manuscript received 21st June, 1967.

THE EFFECT OF PRIOR RECALL ON MULTIPLE RESPONSE RECOGNITION*

BY

JOHN BROWN† and D. W. PACKHAM

From the Department of Psychology, Birkbeck College

Certain investigators have found that recognition is impaired when a recall test is interpolated during the retention interval. One possible explanation of this finding is that interpolated recall leads subjects to employ a more stringent recognition criterion. In the two experiments reported here, the influence of the recognition criterion was eliminated by using a multiple-response test requiring subjects to rank a recognition list consisting of old and new items. Nevertheless recall impaired subsequent recognition in both experiments, the effect being most marked for lowly ranked items. The recognition test in the first experiment was carried out in two stages. This made possible a direct examination of whether recall has an effect on the recognition criterion. No evidence for such an effect was obtained. Other ways in which recall may affect the recognition criterion are discussed.

Introduction

Belbin (1950) and Kay and Skemp (1956) found that recall shortly after presentation interfered with recognition of the stimulus material some minutes later. This is contrary to the following simple analysis of what ought to happen: recall should strengthen the retention of items recalled and leave unaffected the retention of items not recalled so that, on average, performance in a subsequent recognition test will be *facilitated*. Now Hanawalt and Tarr (1961) did obtain some evidence for a facilitatory effect. However, if all these experiments are reliable, we have to conclude that facilitation or interference can be obtained, depending on factors yet to be determined.

One important factor is likely to be the nature of the recognition test. Recognition tests can be classified as either free or forced choice. In a forced-choice test the subject must choose a specified number of the items presented. For example, in the familiar multi-choice test, only one of the presented items is correct and he is normally required to choose just one item. In a common form of free test, a number of new items and a number of old items are presented in random order and the subject is required to pick out those he recognizes, i.e. he is free to choose as many or as few items as he pleases. (In a relatively free test, he would be given upper and lower limits to the number of choices.) The distinction between free and forced-choice tests is important for the following reason. In a free test, performance will be affected by the subject's criterion for recognition, that is, his willingness to pick out items about which he is unsure. In a forced-choice test, on the other hand, the number of items he chooses is prescribed and his task is simply to pick out the most probable items.

Both Belbin and Kay and Skemp used free choice tests. Belbin asked her subjects to accept or reject a picture as identical to a picture seen previously. Here the subject was free to choose either all the presented material or none of it. Kay and Skemp also used this form of recognition test and followed it by a second free test in which the subjects went through a list of items from the picture and picked out those they recognized. In contrast, Hanawalt and Tarr used forced-choice recognition: recognition of each item was tested by a multi-choice test. Thus the possibility arises that recall interferes with subsequent recognition only when a recognition test is used which permits the subject's recognition criterion to influence how many items he chooses. Indeed, under the same set of conditions, recall might interfere with subsequent free recognition yet facilitate subsequent forced recognition. Experiment I below is concerned with this possibility and examines the role of the recognition criterion.

* Experiment I was reported at the Oxford Meeting of the Experimental Psychology Society in 1965. Both experiments were conducted at Birkbeck College, University of London.

† Now at Department of Psychology, University of Newcastle upon Tyne.

Multiple response recognition tests. Both the experiments to be described use a multiple response recognition test (Brown, 1965). In such a test the subject is presented with a list consisting of k original items and of n new items ("distractors"), in random order His task, implicitly or explicitly, is to rank all the items so that the item most likely to be old is put first and the item least likely to be old is put last. Overall performance in the test is evaluated in terms of a measure based on the average rank assigned to the correct items. An important feature of the measure is that it is uninfluenced by the criterion used for recognition, since the subject is required to rank *all* the items. In this respect it is like a forced-choice recognition test but it should be noted that the subject is not asked to accept or reject *any* of the items. Another feature of the measure is that it reflects the degree of discrimination shown in relation to each old item, since each item is assigned a rank. The actual measure used is called $\bar{A}$ and can be immediately derived from the average rank. However, its character is more clearly revealed by the following account. The number of distractors written down after the first old item is counted and divided by the total number of distractors n. This is repeated for each of the remaining $(k - 1)$ old items. The average of the scores so obtained is an estimate of the average probability of rejecting a distractor and is called $\bar{R}$. If all the items are written down in random order, the expected value of $\bar{R}$ is $\frac{1}{2}$. Accordingly the preferred measure is $\bar{A}$, where $\bar{A} = 2\bar{R} - 1$. This has an expected value of 0 when performance is at the chance level and of unity when recognition is perfect.

Experiment I

This experiment was designed to investigate (1) whether recall influences the criterion adopted for subsequent recognition and (2) how recall influences recognition performance when the possible influence of the recognition criterion is eliminated. A two-stage recognition test was employed. In Stage I, the task was to write down the items recognized from a list of new and old items. In Stage II, the task was to write down the remaining items so that the items most likely to be old were put first and the items least likely to be old were put last. Provided the subject's writing down order reflects his ranking of the items in both stages, the following analyses can be made. First, the proportions of subjects choosing correctly at the end of Stage I and at the beginning of Stage II can be examined to see whether prior recall affects the recognition criterion. Second, the two stages taken together can be regarded as a multiple response recognition test and a measure of recognition performance can then be obtained which is independent of the recognition criterion.

Method

Subjects. The experiment was conducted at the beginning of a first year laboratory class. During the introduction the 40 subjects were informed that they were to participate in a group experiment involving colour association. They were then assigned to two separate rooms on a random basis so that there were 10 experimental and 10 control subjects in each room.

Stimulus material. The following materials were prepared prior to the experiment:

(i) Foolscap sheets with the colour red, green, blue, brown, yellow, grey, black, white in this order in each of 28 rows.

(ii) Twenty-eight names of household items recorded on magnetic tape at 3-sec. intervals. These names had been selected at random from a pool of 60 names.

(iii) Instructions defining the experimental and control groups, in sealed envelopes.

(iv) A questionnaire related to the laboratory course.

(v) Recognition lists, which contained 10 only of the 28 recorded words, together with 14 new words from the pool. The lists were set out in two columns of 12 words and a different random order was used in the two rooms.

Procedure. The experiment was in four phases.

Phase 1. The sheets of colour names and envelopes were handed out. Subjects were told that their initial task would involve underlining one of the eight colours they thought appropriate each time a word was read out. The tape-recording of the 28 household items was then played.

Phase 2. Next, subjects were asked to turn over the sheet of colour names and to open the envelopes in front of them. Experimental subjects found an instruction to recall

the items they had just heard: the control subjects found an instruction to write down Christian names of the opposite sex, indicating their preference on a 5-point scale. Phase 2 lasted 5 min.

Phase 3. All subjects completed the questionnaire. This extended the retention interval before recognition by 10 min.

Phase 4. The recognition test consisted of two distinct tasks, viz. Stage 1, to write down only the words recognized as being among those read out previously; Stage 2, to write down all the remaining words from the recognition list, putting first the item most likely to be from the original list, the next most likely item second and so on. The recognition lists were folded and sealed so that the second set of instructions were not seen until the first set had been carried out. Once the subject had turned over to Stage 2, he was not permitted to make any alterations to Stage 1.

Experiment I Results

The first question examined was whether the order in which the items occurred in the recognition list had much effect on the order in which the subjects wrote them down. The rank correlation coefficient between list order in columns and R-values for individual words was only about 0·1. Although there is some difficulty in interpreting this result (because individual subjects may well have scanned across the rows rather than down the columns of the recognition list) it seems fair to assume that list order was of minor importance.

Next it was assumed that subjects had written down the items in the order of their subjective probability (probability of correctness) even in Stage I of the test. Since most of the items written down in Stage I were correct, this assumption is not critical for the interpretation of the results.

Recognition scores for the experimental and control groups were then computed, using $\bar{A}$ as the measure. Mean values of $\bar{A}$ were 0·734 and 0·832 respectively but the difference was not statistically significant ($t = 1{\cdot}22$, $d.f. = 38$).

Comparison of the recall scores showed that out of a total of 348 items written down by the experimental group, 112 (mean 5·6 per subject) items were old items and included in the recognition test, a further 197 were correctly recalled but not on the recognition list and the remaining 39 items were errors. A more detailed analysis of the errors indicated that only 12 items occurred as distractors in the subsequent recognition test and that these items were not consistently chosen before other distractors.

A plausible hypothesis is that any adverse effect of recall on subsequent recognition will be stronger for less well-learned items (Kay and Skemp, 1956). Accordingly values of $\bar{A}$ were also computed for the last four correct items written down by each subject: for the experimental group 0·44 of these items had been recalled. The values were 0·500 and 0·688 for the experimental and control groups respectively and the difference was significant at the 0·05 level ($t = 2{\cdot}38$, $d.f. = 38$). For the experimental group the mean value of $\bar{A}$ in the test as a whole was found to be 0·971 for recalled items: it was 0·736 for items not recalled.

No effect of prior recall on the criterion for recognition was apparent. In both experimental and control groups, 15 out of 20 subjects selected a correct (old) item as their last attempt in Stage I and three out of 20 subjects selected a correct item as their first attempt in Stage II. Strictly speaking, these proportions should be related to the number of correct items remaining at each attempt, but these were similar for the two groups. These results suggest that there is no large effect of recall on the recognition criterion, although many more subjects would have to be tested in order to detect a small effect: an N of 20 is comparatively small for estimating a proportion.

Experiment II

The results of the previous experiment were discouraging for the hypothesis that recall affects the criterion subsequently used for recognition, since a null effect was obtained. On the other hand, they did suggest that recall can interfere with the efficiency of recognition even when the possible influence of the criterion is eliminated (much to our surprise at the time). The aim of this second experiment, therefore, was to see whether more convincing evidence for this interference could be obtained. To this end, several changes were introduced. Some were aimed simply at reducing the overall level of performance in recall and recognition. The main change was a new procedure (devised by D.P.) for the recognition test. Each subject received a pack of playing cards printed with old and

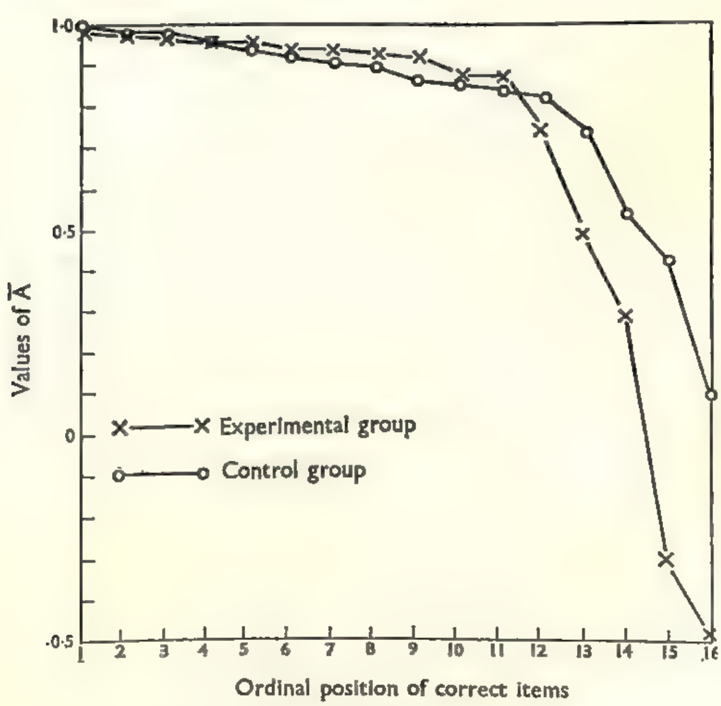
new items. His task was to order the cards (along the old-new dimension of probability) first by sorting them into four piles and then by changing the order within each pile where necessary. One advantage of this procedure is that it enables him to make crude judgements before he attempts finer ones. Another is that it helps to reduce the possible effects of the order in which the items are initially presented in the test. Not only can each subject receive his pack in a different random order but in addition it is easy for him to change his mind concerning the appropriate positions for particular items.

Method

Subjects. Thirty-two first-year psychology students were randomly assigned to experimental and control groups of equal size and were then tested simultaneously in the same room.

Procedure. The experiment was conducted in the same sequence of four phases as Experiment I. The following changes were introduced: (*a*) In order to reduce the overall level of performance, the colour association task was replaced by instructions to categorize 32 words selected from the same pool into metallic and non-metallic objects and the rate of presentation was increased to a speed of 1 word every 2 sec. (*b*) Booklets were substituted for separate instructions in sealed envelopes and the questionnaire was omitted. This shortened the interval between recall and recognition from 10 to 5 min. This interval was occupied by collecting response booklets, by subjects reading the instructions for the multiple-response test and in answering clarificatory questions on the procedure. (*c*) The recognition lists were replaced by packs of 32 printed playing cards. In each pack the first and last eight words which the subjects had categorized were discarded and instead 16 cards printed with new words were selected. The packs were shuffled so as to produce a different random order for each subject. His task was to sort the cards into four piles designated definitely original, probably original, probably not original and definitely not original, and then to put the words in each pile in order of probability. The final order of the pack thus ranged from the most familiar to the most unfamiliar word for each subject.

FIGURE 1.



Value of $\bar{A}$ for successive correct items.

Experiment II Results

The results were broadly similar to Experiment I but the inferiority of the experimental group on the recognition test was more clear-cut. Thus the groups differed significantly even on the overall values of $\bar{A}$($t = 2{\cdot}29$, $d.f. = 30$, $p < 0{\cdot}05$): the actual values were 0·784 and 0·688 for the control and experimental groups respectively. When the groups were compared on the four correct words assigned the lowest ranks, the difference was significant at the 0·01 level ($t = 3{\cdot}87$, $d.f. = 30$). The reason for the higher value of t was that the groups only differed on the terminal words (substantially, this was also the case in Experiment I). This is shown in Figure 1 in which the values of $\bar{A}$ for successive correct items are plotted. It will be seen that the final values of $\bar{A}$ for the experimental group are negative. This implies that recognition performance was at chance level for the last few items: when ranks are assigned at random, some positive and some negative values of $\bar{A}$ tend to result, although the mean expected value will be zero.

The fact that the groups differed only on terminal words suggests that the adverse effect of recall is confined to poorly learned items. The following analysis was performed to check on this possibility: the values of $\bar{A}$ for individual words was calculated separately for the experimental and control groups. (This involved pooling across ordinal positions: the results shown in Figure 1 involved pooling across words). The rank correlation between the value for the control group and the difference between the values for the control and experimental groups was then calculated but was only 0·087. Thus there was no evidence that the words which were on average poorly learned by the control group were most likely to suffer from the effect of interpolated recall. The values of $\bar{A}$ for the control group on individual words ranged from 0·656 to 0·936.

For the experimental group, mean recall was 17·75 of the 32 words presented and was 8·62 of the 16 words included in the recognition test. Intrusions were 8·5 per cent of the total number of recall attempts: about half of the words occurring as intrusions were distractors in the recognition test. The mean values of $\bar{A}$ for recalled and non-recalled words were 0·940 and 0·358 respectively.

Discussion

Recall interfered with subsequent recognition in both experiments. Various theories to account for such interference will now be considered.

Criterion shift. As was pointed out in the introduction, recall might shift the criterion used for recognition in a free recognition test so that fewer items were recognized. This theory will not account for our results since a measure of recognition ($\bar{A}$) was used which is independent of the criterion. Moreover, no evidence was obtained in Experiment I to suggest that the criterion *is* affected by prior recall.

Threshold shift. Kay and Skemp (1956) advanced the theory that the juxtaposition of strong and weak associations raises the threshold for the latter. This could account for the adverse effect of prior recall if recalling some items increases the relative strength of the associations concerned. The theory is identical with the first if a rise in threshold merely means that the criterion for recognition has been raised. Other interpretations are possible, however. For example, suppose an association below threshold is held to be incapable of affecting performance in a recognition test. Then a change of threshold could affect the rank assigned to an item. A fall in the threshold could bring a marginal association above threshold and vice versa. Thus at least one form of threshold shift theory might explain our results.

Erroneous learning. There are three ways in which erroneous learning might influence performance in a multiple-response test. (*a*) The overt intrusion of an item which is a distractor in the recognition test could lead to its erroneous learning as a correct item. As a result, it might be assigned a higher rank in the recognition test and this would tend to lower the value of $\bar{A}$. For the subjects who produced distractors during recall in Experiment II, the mean rank assigned to these distractors was 9·1 as compared with 22·7 for the remaining distractors. However, no relation of cause and effect can be assumed, since the more plausible distractors are the ones most likely to intrude in recall. (*b*) The implicit intrusion of a distractor during recall could have a similar effect. If items merely come into the subject's mind during recall, this may give them a flavour of recency which misleads him during the subsequent recognition test. (*c*) The overt or implicit intrusion of an item similar to a distractor could also disturb the subjects' performance in the test through generalization. (It is of interest to note that Belbin, 1950, attributed her results to a form of erroneous learning. She suggested that rejection of a picture in a recognition test could result from previous erroneous recall of some feature of the picture.)

Erroneous unlearning. The recall task involved free recall, i.e. subjects could write down the items in any order. If a subject thinks of a correct item during free recall but erroneously rejects it, this might merely lead to unlearning of the item. This would tend to lower $\bar{A}$. Unlearning is now a well-established phenomenon and occurs in a wide range of situations including free recall learning (Asch and Ebenholtz, 1962).

In evaluating any of these theories, it is important to bear in mind that $\bar{A}$ will be lowered only if the average rank assigned to correct items is lowered. Moreover, the factor or factors leading to interference may be offset, or more than offset, by a tendency for recall to strengthen the learning of the items recalled. The effect of such strengthening is most likely to be dominant when recall takes place while retention is still high and when the recognition test is delayed. Indeed, Hanawalt and Tarr (1961) found that the facilitatory effect of recall on recognition in their experiment was much more marked with a recognition test delayed for 24 hr., than with a test immediately following recall.

The fact that Hanawalt and Tarr obtained a facilitatory effect even in the immediate test calls for comment. One possibility is that their results were determined by the particular construction of their recognition tests. Each correct word was presented with four distractors, one of which was a synonym or an opposite of the word. Any subject who detected this feature will have had a 1 in 2 instead of a 1 in 5 chance of a correct response when guessing. And recall subjects may have been more prone to detect it, perhaps because they tended to be very confident about the words they had successfully recalled. A subsidiary experiment of Hanawalt and Tarr (1961) is worth mentioning. It was based on an experiment of Postman, Jenkins and Postman (1948) and produced results unlike those of their main experiment. Postman *et al.* had found that multi-choice recognition of nonsense syllables was impaired by prior recall but the comparison (which was incidental to the main purpose of the experiment) was between immediate recall and recognition delayed by recall. Hanawalt and Tarr held the retention interval before recognition constant at 10 min. and obtained no significant difference between conditions. However, it is noteworthy that their actual results were close to those of the original experiment: their recognition means for the recall and no-recall conditions were 24·69 and 26·94 respectively as compared with 24·23 and 27·66 in the original experiment.

References

Asch, S. E., and Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *J. Psychol.*, **54**, 3–31.

Belbin, E. (1950). The influence of interpolated recall upon recognition. *Quart. J. exp. Psychol.*, **2**, 163–9.

Brown, J. (1965). Multiple response evaluation of discrimination. *Brit. J. math. statist. Psychol.*, **18**, 125–37.

Hanawalt, N. G., and Tarr, A. G. (1961). The effect of recall upon recognition. *J. exp. Psychol.*, **62**, 361–9.

Kay, H., and Skemp, R. (1956). Different thresholds for recognition—further experiments on interpolated recall and recognition. *Quart. J. exp. Pxychol.*, **8**, 153–62.

Postman, L., Jenkins, W. O., and Postman, D. L. (1948). An experimental comparison of active recall and recognition. *Amer. J. Psychol.*, **61**, 511–9.

Manuscript received 19th April, 1967.

COMMENTS ON "SELECTIVE ATTENTION: PERCEPTION OR RESPONSE?"

BY

J. A. DEUTSCH and DIANA DEUTSCH

From the Department of Psychology, University of California, San Diego

AND

BY

P. H. LINDSAY

From the Department of Psychology, University of California, San Diego

REPLY

BY

ANNE M. TREISMAN

From the Institute of Experimental Psychology, University of Oxford

We cannot understand why Treisman and Geffen (1967) think their experiment argues against our theory (Deutsch and Deutsch, 1963). Briefly, Treisman and Geffen ask subjects to repeat and tap to certain words in one message, played to one ear, and only tap to such words when they occur in another message played to the other ear. They find that subjects neglect the words to which they only have to tap. According to our theory, stimuli with a greater weighting of importance inhibit certain outputs (such as storage, motor response) of the structures processing stimuli with a lesser weighting of importance. Now it seems to be clear that Treisman and Geffen have by their instructions (to tap and repeat one set of words and only to tap to another set of words) produced a situation in which one set of stimuli is given a larger weighting of importance than the other. It is therefore not surprising on our theory that the less important set is almost disregarded. It is instructive here to consider Lawson's (1966) very similar experiment. In this experiment the signals to which the subject has to tap do not also have to be repeated if they occur in the message which is being shadowed. (These signals are non-verbal.) Lawson's results are almost the opposite of Treisman and Geffen's, as would be expected from our theory. Treisman and Geffen have some difficulty in explaining the discrepancy. "It seems that analysis of simple physical signals precedes both the selective filter and the analysis of verbal content in the perceptual sequence, that the bottle-neck in attention arises chiefly in speech recognition where of course the information load is usually much higher. To confirm the belief that the verbal content of the secondary message was not being analysed, we find no evidence whatever of interference from secondary target words when these received no tapping response." (We quote the last sentence as just one example of the fact that Treisman and Geffen have failed to understand our theory. It is one of the major points of this theory to explain why "secondary" messages do not cause interference with the "primary" message while they are being analysed.) To return now to the subject of Lawson's experiments, we would suggest that the outcome of such experiments would be the same if instead of signals, words were used in Lawson's paradigm. These words should occur on both channels and should be distinguishable by another speaking voice. The subject should be asked to respond to, but not to repeat such words. To make sure the subject is not simply responding to differences in timbre, pitch, etc., the target words should be interspersed with other words. Treisman and Geffen could not then postulate differences in information load to explain an unfavourable result.

Finally we would like to make a comment on Treisman's own suggested amendment of Broadbent's theory. "If the filter reduced the signal-to-noise ratio of unattended messages rather than blocking them completely, words which were highly important or relevant to the subject might still be perceived despite this attenuation, provided that the criteria for detecting them were sufficiently low. This would have the biological advantage that the unattended messages could be monitored for any important signals without at the same time much increasing the load on the limited capacity available for speech recognition." (Treisman and Geffen, 1967). It would seem to us that Treisman's suggestion of attenuation would have quite the opposite effect. It seems evident that on the whole, a signal-recognizing (as distinct from signal-transmitting) system would be much more disrupted or taxed by having to recognize signals which were incomplete or noisy, than if such signals were clear. The "load" would clearly be increased over the case where no attenuation occurred. Similarly, a signal recognizing system would have to increase the amount of processing when it had to distinguish between signals some of which were incomplete or noisy. The introduction of corrupted messages could certainly never reduce the load on the system as compared with the case where such messages were not degraded. Such degradation would simply reduce the efficiency of decision made by a system. Treisman's proposed amendment to Broadbent's theory makes matters worse.

J. A. DEUTSCH AND D. DEUTSCH.

REFERENCES

DEUTSCH, J. A., and DEUTSCH, D. (1963). Attention: some theoretical considerations *Psychol. Rev.*, **70**, 80–90.

LAWSON, E. A. (1966). Decisions concerning the rejected channel. *Quart. J. exp. Psychol.*, **18**, 260–5.

TREISMAN, A., and GEFFEN, G. (1967). Selective attention: perception or response? *Quart. J. exp. Psychol.*, **19**, 1–17.

Manuscript received 24th July, 1967.

A recent report by Treisman and Geffen (1967) suggests there is some ambiguity in the interpretation of a theory of attention proposed by Deutsch and Deutsch (1963). Perhaps one reader's view of the controversy might clarify some of the issues. The theory proposed by the Deutschs does not question the fact that the processing capacity of the nervous system is limited—after all it is a finite system; nor do they quarrel with the proposition that at some stage during the analysis of incoming messages there occurs a "single channel mechanism" which gives preferential treatment to some messages at the expense of others. Their theory is concerned with the question of where the single channel process arises. It is motivated by the consistent finding that complex stimuli, such as particularly important words, when they occur on a secondary channel supposedly being ignored do succeed in diverting a subject's attention from a primary task. They argue that these findings indicate the single channel process must occur after the messages have been fully analysed by the perceptual system and therefore impaired performance found under multitask conditions might be due to limitations in the storage of messages and responsiveness to them rather than their initial perceptual analysis. Since most tasks used to study attention involve a large memory component, there are no convincing grounds for attributing the observed decrements in performance to processes operating during the perception of the stimulus.

The main thrust of Treisman and Geffen's experiment appears to be that memory is not involved in their particular task and consequently decrements in performance cannot be associated with memory limitations. The point is debatable. On the average, several items intervene between the presentation of a target word and the subject's response. Interference effects on both messages often span a long string of items. These time periods cannot be accounted for in terms of simple reaction times and suggest that the storage of information is a necessary factor in the performance of Treisman and Geffen's experimental tasks.

Treisman and Geffen also show that, though the analysis is admittedly hazardous due to the small number of observations and the assumptions needed to determine the false alarm rate, the d′ statistic is lower for unshadowed messages than for shadowed messages. They imply that this difference in d′ indicates information is being lost during the perceptual analysis of the unshadowed material. While the reduction in d′ does indicate that information is being lost somewhere between the presentation of the signal and the subjects' response, it says nothing about where the loss occurs. It could be reasonably argued that

because of its relative unimportance, as emphasized by the instructions, the output of the perceptual analysis of the unshadowed message is only briefly and infrequently monitored perhaps at opportune times during the shadowed message. This would reduce the "hit rate" for target words in the unshadowed message. Since, according to the Deutschs' theory, performance on the dominant task prevents information in the secondary message from being transferred into a long-term storage, the subjects will often be faced with trying to decide whether a target word has occurred in the unshadowed message on the basis of a fast fading (STM) memory trace of the recently presented words. This would reduce the possible precision of discrimination and therefore measured detectability.

Treisman and Geffen point out the apparent contradiction of their results and the results of a study by Lawson (1966) who found no impairment when subjects monitored several non-verbal signals at the same time. The explanation offered seems to rely on an intuitive appeal to the relative processing capacity involved in the analysis of simple physical signals as opposed to that involved in the analysis of verbal messages. Processing capacity, however, is neither explicity defined or measured. Moreover it is probable that the information load on a processor is not simply related to the continuum of task "complexity." M. M. Taylor, S. M. Forbes, and I (1967) have found that even with standard psychophysical discrimination tasks such as the discrimination of the pitch or the intensity of a pure tone, the requirement of performing several such tasks at the same time substantially reduces the precision of discrimination for each task. Processing load may be related to the difficulty in discriminating between two signals rather than the complexity of the analytic operations assumed to be performed. These findings, however, again raise the question of the degree to which limitations in the storage of information would plausibly account for the impairment found with simple discrimination tasks.

Finally some clarification of the notion of attenuation would be useful. In particular it is not clear how the attenuation process affects the discrimination and identification of irrelevant messages. Simply attenuating a secondary message would not alter the signal to noise ratio and therefore the precision of discrimination would be unaffected. It does not seem logical to assume that just the "signal" components of an incoming message would be attenuated since this requires prerecognition of the message. A reduced signal to noise ratio would result if the message was attenuated prior to the addition of neural noise but an explicit system of this type is difficult to conceptualize and would hardly be an economical method for identifying and rejecting irrelevant messages.

P. H. Lindsay.

References

Deutsch, J. A., and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.*, **70**, 80–90.

Taylor, M. M., Lindsay, P. H., and Forbes, S. M. (1967). Quantification of shared capacity processing in auditory and visual discrimination. In Sanders, A. F. (Ed.). *Attention and Performance*. Amsterdam: North-Holland.

Treisman, A., and Geffen, G. (1967). Selective attention: perception or response? *Quart. J. exp. Psychol.*, **19**, 1–17.

Manuscript received 24th July, 1967.

Deutsch and Deutsch and Lindsay have raised some criticisms of the paper "Selective attention: perception or response?" by Treisman and Geffen (1967). Briefly, the argument of this paper was as follows: in tasks requiring that attention be selectively directed to one of two competing speech messages, subjects can report little or nothing of the verbal content of the unattended message. This might result from a limit either in perception (as suggested by Broadbent, 1958, in his "filter" model) or in response mechanisms (as proposed by Deutsch and Deutsch, 1963). An experiment was performed to discover which of these stages constituted the main bottleneck. Subjects were presented with two competing speech messages given simultaneously, one in each ear, and were asked to make two responses: the primary response was to repeat back the attended message; the secondary response was to tap whenever a specified target word was heard on either ear. If the main limit was reduced perception of the secondary message, the tapping response to the primary message should be much more efficient than that to the secondary message. But if the main limit lay in the organization of two simultaneous responses, the secondary, tapping response should be equally inefficient whichever message contained the target word, since this word

was the same for the two messages and so should be equally inhibited by the primary repeating response. We argued that our results strongly favoured the limit being perceptual, since subjects tapped to 87 per cent. of primary message targets and to only 8 per cent. presented in the secondary message.

Deutsch and Deutsch and Lindsay raise a number of different theoretical points of some importance. I believe most of the criticisms are based on misunderstandings, so it may be useful for me briefly to clarify our suggestions on each point in turn.

(1) *The effect of importance.* We agree, of course, with Deutsch and Deutsch that the direction of attention can be biased by the relative importance of stimuli. Our interest lay in discovering whether the bias chiefly affects perceptual processing or response mechanisms in the selective listening task. Our instructions were intended to stress the importance of the verbal response to the primary, "attended" message, but to give equal though lesser weight to the tapping responses to targets in both primary and secondary messages. Deutsch and Deutsch suggest that the requirement to repeat as well as tap to the primary targets adds to their importance. This is possible, although our subjects actually seemed more concerned about failures and successes with secondary than primary targets. I agree that it would be useful to repeat the experiment without this requirement and an attempt to do this is referred to below. However it seemed to us that, logically, if a subject's response and memory capacity will support both repeating a word and tapping to a target in 87 per cent. of cases when both are in the primary message, it should support the same level of response when the target is in the secondary message, provided that perceptual analysis is complete in all cases, as Deutsch and Deutsch suggest. It is surprising, then, that only 8 per cent. correct responses were made to secondary targets. Moreover the verbal context of the target words and their homophones affected primary and secondary messages very differently, which is inexplicable if there were no differences in the perceptual analysis of the two messages.

(2) *Interference from secondary targets.* We interpreted the absence of any interference from secondary targets when not tapped to as evidence against full perceptual analysis. Deutsch and Deutsch consider this misinterprets their theory, which, they say, specifically accounts for the lack of such interference. However, our interpretation of their theory was based on their 1963 paper, which quotes Peters's (1954) report of interference by a secondary message as evidence supporting their theory. Moreover we also found that the interference from secondary targets which *were* detected was greater than that from primary targets. We argued that, since the two tapping responses were identical, they should not differ in response or memory load and so detection of the secondary targets must have interfered with *perception* of the primary message.

(3) *The role of memory.* Lindsay argues that our task involves short-term memory. This is true, although it is much less dependent on memory than most tasks which have tested selective attention. The mean response lag for both primary and secondary targets was only 1·2 sec. (3 items) which does not allow much time for decay in a post-perceptual store. The equality of the reaction time to those primary and secondary targets which were detected argues against the intermittent, often delayed monitoring of a stored trace of secondary targets. The time span of interference effects need have nothing to do with storage time; it could equally well be due to losing track, delay in picking up the correct message and so on, even after the interfering item has been forgotten.

(4) *Lawson's results.* Deutsch and Deutsch claim that we have some difficulty in explaining the discrepancy between Lawson's results (1966) with tones as targets and ours with words. So far from this being true, we actually predicted the difference on the basis of Broadbent's theory (see Treisman, 1964*a*), since differences in physical characteristics must be analysed before the "filter" selects the attended message. Evidence supporting this claim is that (*a*) these physical characteristics can all be used as a basis for selection of a message to be attended to (unlike verbal differences such as a change of language); (*b*) they can be reported even for unattended messages; and (*c*) variations in these features can interfere with attention to the selected message (again unlike changes in verbal features of a message). This conclusion is not intuitive, as Lindsay supposes, but based on experimental evidence (Cherry, 1953; Treisman, 1964*b* and *c*). Lindsay also argues that processing capacity is not simply related to stimulus complexity and in support refers to the finding that simultaneous psychophysical discrimination tasks may interfere with one another. But we have not argued, nor would it seem plausible, that there is only one way in which tasks can interfere with one another. The fact that simple discriminations may compete

in threshold experiments does not imply that complexity of perceptual analysis is irrelevant to the difficulty of the task in the selective listening situation we were investigating.

(5) *A test of the theories.* Deutsch and Deutsch's theory is difficult to test, since it normally precludes subjects showing either by response or recall that they have analysed the content of secondary messages. However they do propose a crucial test: our experiment should be repeated with no verbal response made to primary targets, which, I agree, is a necessary control. They also require that the target words should be in a different voice from the rest of the passage, though why this should be necessary to their prediction is not clear. In any case I would also predict that targets in a different voice *would* be detected in the secondary message (for the same reasons as Lawson's tones), and this would possibly be found even if control, non-target words were also used. A difference in voice is one of the physical characteristics which appears to be analysed before the filter, and, if the task makes this cue available, subjects may use it to reduce the ensemble of possible secondary targets which they must monitor to manageable proportions.

A short experiment (which will be reported more fully elsewhere) was run to test the predictions. Sixteen lists of 16 pairs of digits were recorded in a man's voice at 1·8 pairs a sec. The stimuli were recorded on digital tape, equated in length at 250 millisec. by computer compression or expansion and exactly synchronized. At different positions in each list one digit was replaced by a letter. In half the lists this was in the same man's voice as the digits and in half in a woman's voice. Eight lists, randomly chosen, had the letter on one track and eight on the other. The seven subjects were asked to attend to and repeat back the digits on the right ear, but to stop repeating and tap at once if they heard a letter on either ear. This was intended to avoid both response competition and memory limitations. Subjects heard the 16 lists twice through; the primary lists on the first run were the secondary ones on the second run. Subjects were always told which voice would speak the letter and on one run they were told what the letter would be, but they never knew in which ear or which list position it would occur.

The results were as follows: primary message, same voice—71 per cent. correct; primary message, different voice—97 per cent.; secondary message same voice—28 per cent.; secondary message different voice—97 per cent. Subjects found shadowing this computer-synchronized material much more difficult than the prose in the previous experiment and only repeated 74 per cent. of correct items compared to 93 per cent. in the previous experiment. They also frequently switched to the other channel, repeating 13 per cent. of items from the secondary message; this never occurred in the previous experiment. This inability to select the message on the correct ear consistently was most probably due to the accurate synchronization of the digit pairs. If we assume that for at least 13 per cent. of the time the secondary message was actually functioning as the primary message, the true detection rates become at least 79 per cent. for primary targets and at most 20 per cent. for secondary ones. (This assumes that during the 13 per cent. omissions subjects were not switching their attention at all.)

The result confirms our prediction that, when subjects cannot select before the filter on the basis of voice quality, they are unable to detect the majority of secondary targets. Our previous result was not therefore due to the requirement to repeat as well as tap to primary targets, and the discrepancy from Lawson's results with tones is not explained by this suggested difference in relative "importance." With targets in a different voice, which made preselection possible, subjects detected almost all the targets in both ears. This argues against an inherent response limit and supports the theory of a perceptual limit arising chiefly at the stage where the verbal content is identified.

(6) *Selection by reduction in signal-to-noise ratio of secondary messages.* First, we did not wish to imply, as Lindsay suggests, that the possible difference in the d′ statistic was evidence for perceptual selection. Our point was that—given that the rest of our results showed a perceptual limit—then the form it took appeared to be a reduction in S/N ratio for secondary messages (although the evidence is still admittedly tentative). Both Deutsch and Deutsch and Lindsay raise a more general point. They doubt that any economy in perceptual analysis could be gained from a reduction in S/N ratio, and claim that this would have the opposite effect of making rejection more laborious. But this criticism seems to be based on unnecessary assumptions. Talk of disruption by noise, as though detection were obligatory and noise *had* to be combated, seems misconceived. The assumption we made was that words presented to either channel are exposed to a series of detection processes which can either detect or fail to detect, and no special disruption is entailed by either

outcome. A possible system for speech recognition consists of a hierarchical arrangement of tests for the critical features distinguishing phonemes, words or phrases. If unshadowed words, with their low S/N ratio, can be discarded at the earliest stages when they fail the most general tests, this would eliminate any interference at subsequent stages of speech analysis. This rejection of irrelevant words would be unfortunate in the case of very important items (own names, target words etc.) but in these cases the effect of the lowered S/N ratio could be combated to some extent by maintaining the criteria for their detection especially low. Thus, although detection of these important words in the secondary message would be less frequent than in the primary message (as it is in fact), there would still be considerable economy, in that most of the secondary stimuli would not be analysed in any detail. There is some evidence that a noisy message does in fact interfere less with perception of a competing message than one that is clear and easy to perceive (Treisman, 1964*b*).

Finally Lindsay would like some clarification of the notion of attenuation, or reduction in effective S/N ratio. We deliberately left this suggestion as general as possible, since behavioural evidence is unlikely to support one precise alternative against another. However, his suggestion of attenuation followed by the addition of noise was one possibility we had in mind; in a multi-stage perceptual system this process does not seem to us either implausible or "difficult to conceptualize."

ANNE M. TREISMAN.

REFERENCES

BROADBENT, D. E. (1958). *Perception and Communication*, London: Pergamon.

CHERRY, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *J. acoust. Soc. Amer.*, **25**, 975–9.

DEUTSCH, J. A., and DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.*, **70**, 80–90.

LAWSON, E. A. (1966). Decisions concerning the rejected channel. *Quart. J. exp. Psychol.*, **18**, 260–5.

PETERS, R. W. (1954). Competing messages: The effect of interfering messages upon the reception of primary messages. *USN Sch. Aviat. Med. res. Rep. Project No. NMOO1* 064.01.27.

TREISMAN, A. M. (1964*a*). Selective attention in man. *Brit. Med. Bull.*, **20**, 12–6.

TREISMAN, A. M. (1964*b*). The effect of irrelevant material on the efficiency of selective listening. *Amer. J. Psychol.*, **77**, 533–46.

TREISMAN, A. M. (1964*c*). Verbal cues, language and meaning in selective attention. *Amer. J. Psychol.*, **77**, 206–19.

TREISMAN, A. and GEFFEN, G. (1967). Selective attention: perception or response? *Quart. J. exp. Psychol.*, **19**, 1–18.

Manuscript received 4th October, 1967.

THE EXPERIMENTAL PSYCHOLOGY GROUP, 1946-1958*

The idea of the Experimental Psychology Group goes back to the last year of the war when some of us were starting to think about the shape of post-war psychology and our own relations to it. On the one hand, we hoped that the gains the war had brought to psychology in the shape of realistic experiment and close relationship with other scientific disciplines would be preserved and extended. On the other hand, we were worried about the current state of experimental psychology—at that time largely geared to immediate practical ends—and felt that there was real danger of losing sight of basic issues. We were also a bit anxious about the tendencies to professionalism which were so marked in British psychology at that time.

I say "we" because I honestly don't know who first had the idea of the Group. I certainly remember conversations in 1944 with Ralph Pickford and I think also with Kenneth Craik—though of that I am less certain—and in the following year with G. C. Grindley and Carolus Oldfield. At all events, the first real step was taken in June, 1946, when a preliminary meeting was held at Cambridge, in Professor Bartlett's rooms at St. John's College. To this meeting a number of younger people in experimental psychology were invited and the decision to establish a new group formally taken. Seven more people were invited to join us. As regards the name it was agreed that "Experimental" best described our general intention and "Group" its informal and limited character. If the Group is to have a birthday it should certainly be 21st June, 1946.

The first formal meeting of the Experimental Psychology Group was held in the old Library of the Cambridge Psychological Laboratory on 19th October, 1946. According to the Minute Book, those present were Bernard Babington Smith, Julian Blackburn, Derek Russell Davis, George Drew, Hans Eysenck, G. C. Grindley, Edmund Hick, Norman Mackworth, Carolus Oldfield, Ralph Pickford, Boris Semeonoff, John Whitfield and Oliver Zangwill (Convenor). Apologies for absence were received from William Honeyman† and Magdalen Vernon. Pickford was elected President, with myself as Honorary Secretary, and Grindley and Oldfield were elected to the Committee. The scientific part of the meeting consisted of two papers, each of which was followed by considerable discussion. The first was by Grindley and entitled "The transposition of visual patterns" (later published in *Brit. J. Psychol.*, 1947, **37**, 152, in collaboration with Valerie Dees). The second was by Mackworth and entitled "Some internal inhibition effects in a vigilance task" (later published in more extended form in this *Journal*, 1947, **1**, 6). The presentation of these papers, informal and followed by lively discussion, set the pattern of our future meetings.

While we are dwelling on these very early days, I would like to take the opportunity to point out that the Experimental Psychology Group was conceived from the beginning as a small independent discussion group and that its rules specifically excluded its playing any part in activity intended to affect the professional status of its members or psychologists generally. We were thus in no sense in competition with or opposed to the British Psychological Society. At the same time, there can be no doubt that the formation of the Group owed something to misgivings felt by a number of us about certain tendencies current in British psychology at the time. These Oldfield and I expressed freely to Professor A. W. Wolters, at that time a member of the Council of the B.P.S.—and indeed virtually its elder statesman—and we sought an interview to obtain his advice on how we might best present our case for the formation of a new group. Shortly before our meeting Professor Wolters wrote to Oldfield as follows. His letter is dated 1st July, 1946:

> As a preliminary to our talk let me say first that I fully understand the motives which have caused action of this kind. I share the uneasiness you feel as to present tendencies in British psychology. Zangwill heard me mutter something of the kind at Durham, and I spoke more vehemently and bitterly at the end of the meeting.
>
> Further, I warmly appreciate your care for the B.P.S. The Society will always be my first care, and I should like to think of your Group as an organ for its reform. We may be able to discuss some way in which that may develop. For one thing, I wish

* Based on a speech given at a dinner held at The Queen's College, Oxford, on 10th July, 1967, to celebrate the 21st Anniversary of the foundation of the Experimental Psychology Group.

† Dr. Honeyman died on 24th November, 1946, at the age of 34. An obituary notice appeared in the Minute Book.

you younger people would make a determined effort to get a stronger representation in the Council.

Finally, having had to spend much time in sitting on foolish revolts, I should welcome one which I could support, and I think the time is ripe for it.

In a letter dated 29th July, 1946, Professor D. W. Harding, at that time Honorary General Secretary of the British Psychological Society, wrote to me as follows:

The Council of the Society at its last meeting received and considered a statement from Professor Wolters regarding the formation of a group for the study of experimental psychology. The Council was of the opinion that it is open to any members to form such groups for the intensive study of special subjects, and that such activities are to be welcomed.

The Council appreciates the courtesy of members of the experimental group in laying the matter before it, offers its good wishes for their success, and hopes that at a later date the proceedings of the Society may be enriched by the results of their work.

This courteous welcome from the Council of the B.P.S. greatly strengthened our resolve to proceed with the formation of the Group.

In the early days, meetings were held regularly four times a year, as a rule in Oxford, Cambridge or London. By courtesy of Julian Blackburn, the Group met several times at the London School of Economics, where the first annual general meeting was held on 13th December, 1947. Among the more memorable occasions was a meeting in Cambridge in July, 1949, following a Symposium of the Society for Experimental Biology at which papers were read by Karl S. Lashley and Konrad Lorenz. I remember also a meeting held at the Oxford Institute of Experimental Psychology later the same year at which Warren McCulloch electrified us with a paper entitled "How nervous structures have ideas." Then there was a meeting in Cambridge, in July, 1950, at which Sir Frederic Bartlett sketched out his "Programme for Experiments on Thinking" (published in this *Journal*, 1950, **2**, 145), which later evolved into his well-known book on *Thinking: An Experimental and Social Study* (1958). But apart from these star occasions, there were many more modest meetings at which members and invited speakers talked to us about their current work and a tradition grew up of free, outspoken (but almost always good-natured) critical discussion. It also became our custom to hold a party in the course of every major meeting. Both these conventions, I am happy to say, have been carried over into traditions of our present-day Society.

A particularly good suggestion was made by John Whitfield late in 1947 that visiting foreign psychologists should be admitted to temporary membership of the Group. This led to the creation of a new category of Visiting Foreign Membership intended principally for post-doctoral Fellows and others visiting this country for periods of up to one year. The first visiting foreign member was an Indian graduate student, William Adiseshiah, at that time a research student at Cambridge. Roger Russell became a visiting foreign member in 1949 though he soon transferred to ordinary membership on his appointment to the Chair of Psychology at University College, London. Other visiting foreign members were R. R. Blake, C. H. Graham, H. A. Imus, S. Koch, Ardie Lubin, G. H. Mowbray, Dewey Neff and Larry Weiskrantz.

By far the most exciting event in the early history of the E.P.G. was an announcement by the Honorary Secretary that the group had received a benefaction from a member who wished to remain anonymous for the purpose of founding a journal. The announcement was made at a business meeting of the Group at Oxford on 19th July, 1947, and in accordance with the benefactor's wishes the amount of the benefaction was not disclosed. It was hinted, however, that the amount would be sufficient to run a small journal for at least 5 years. After they had recovered from their amazement, the Group accepted this magnificent gift with alacrity and set up a Trust Fund (known as the Experimental Psychology Trust) to administer it. The Group next commissioned the Committee to explore the various possibilities that there might be for starting a journal. Carolus Oldfield, who played a most active part in these explorations, advised the Committee to place the *Journal* in the hands of Messrs. W. Heffer & Sons Ltd. of Cambridge. The first issue of the *Quarterly Journal of Experimental Psychology* appeared under Oldfield's editorship in April, 1948. Most unfortunately, though, Oldfield felt obliged to resign the editorship on account of pressure of other work. He was succeeded in the following year by Derek Russell Davis. Although the scope of the *Journal* was limited by considerations of finance, Russell Davis succeeded admirably in building up a journal that was at the same time of good intellectual standing and economically viable.

It can now be revealed that our benefactor was G. C. Grindley and the amount of his benefaction £1,000. Although Grindley has always—and with characteristic modesty—disclaimed any special responsibility for the *Journal*, there can be no doubt that it was his idealism and generosity that led to its foundation. It should also be said that with his permission the Experimental Psychology Trust was wound up in 1959 and all funds administered by it transferred to the new Experimental Psychology Society.

The question of the future of the Group exercised many of us, and from 1948 onwards our aims, objects and optimal rate of expansion were constantly under discussion. In the first 5 years of its existence the Group had doubled in size, but even so this increase failed to keep pace with the expansion in experimental psychology in the country at large. One proposal was that members of the Group might retire from active membership when they had reached a certain age—not I think precisely specified—and would then constitute a kind of "Upper House," without voting rights. This rather odd proposal was rejected, no doubt rightly. At the fourth annual A.G.M. in December, 1950, a membership limit of 40 for the next 3 years was agreed though it soon became clear that this could be but a temporary expedient. Although some members clearly preferred to continue with a small and specialized group, opinion gradually hardened in favour of the idea of our metamorphosis into a society, which would keep pace with the expansion of experimental psychology in modern Britain and measure up to the responsibilities which we had incurred through our sponsorship of the *Journal*. In formulating and defining its future course, the Group owed much to Carolus Oldfield's persuasive advocacy of a larger and more representative association. And in drawing up a constitution for the new Society, the Group were heavily indebted to their last President, Edmund Hick.

At the twelfth (and last) A.G.M. of the Experimental Psychology Group on 30th December, 1958, membership stood at 64. At this meeting, the Experimental Psychology Group was dissolved and the Experimental Psychology Society, with 74 founder members, duly constituted.

O. L. Zangwill.

APPENDIX

Officers of the Experimental Psychology Group, 1946–58

President:	R. W. Pickford (1946–9)
	G. C. Drew (1950–1)
	M. D. Vernon (1952–3)
	J. W. Whitfield (1954–5)
	R. C. Oldfield (1956–7)
	W. E. Hick (1958)
Hon. Secretary:	O. L. Zangwill (1946–52)
	A. D. Harris (1953–5)
	P. H. R. James (1956–7)
	J. Brown (1958)
Hon. Treasurer:	B. Babington Smith (1948–52)
	W. E. Hick (1953–6)
	J. Szafran (1957–8)

BOOK REVIEWS

Machine Intelligence. Edited by N. L. Collins and D. Michie. Edinburgh: Oliver & Boyd. 1967. Pp. x + 278. 63s.

If there were an exact and universally accepted definition of *intelligence* in humans, or better in animals, then it might be reasonable to enquire about the possibility of intelligence in machines. There are now machines which show adaption and learning; machines which can classify speech and hand-writing; machines which can perform goal-directed motor activity, machines which play board games, such as draughts, very well; machines which can hold a conversation with a man and convince him that he is communicating with another human being. Are all, some, or any combination of these, the signs of intelligence? On viewing the behaviour of these machines for the first time the answer might well be strongly affirmative—on discovering the simple mechanism by which these feats were performed this opinion might well undergo a radical change; nothing is so disheartening as the first view up the conjuror's sleeve.

The real answer to the problem of defining machine intelligence is probably that this is a frontier science, by nature undefinable—when we are having trouble in making a machine act in a manner previously unique to man then it is a problem of *machine intelligence*—when we have solved the problem (and know how we did it) then we have a plain *machine*. Any coherence in the subject-matter of this book lies only in this, that it reports work concerned with attempts to perform tasks which are reasonable for the human being but very difficult (at present) for the machine.

The seventeen papers in this collection were presented at a seminar or *workshop* at Edinburgh in 1965, and represent a major part of British effort in this field. They range from mathematical works on theorem-proving and computer languages, through more heuristic approaches to computer game-playing and problem-solving, to the discussion of machines emulating man's perceptual and verbal skills. Many of the papers suffer a great deal from the lack of any obvious context or purpose—one is not prepared to wade through a mass of mathematical and logical argument without knowing what it achieves and to what this is relevant. Presumably at the seminar the authors filled in the background to their work in informal discussion, or could safely assume that it was known. The book would be very much better if the editors had made this information available to the general reader, introducing each paper and setting it in the context of other work.

The papers with most interest for human psychologists are those by Clowes on picture-processing; by Hill on automatic speech-recognition; by Gregory on seeing-machines; and by Kiss on models of word-storage. This last paper in particular is fascinating in demonstrating how an on-line computer may be used to examine and model higher mental processes simply yet with great power.

Much of the book is dull and plodding, and more forceful editing should have been employed. It is, however, a fair reflection of British work on artificial intelligence, which lags well behind that in the U.S.A. (large machines play a big part in these studies), and contains some interesting pointers to those aspects of intelligent behaviour which are simple to emulate, and those which are not.

BRIAN GAINES.

Mechanisms of Animal Behavior. By P. R. Marler and W. J. Hamilton III. London and New York: Wiley. 1966. Pp. xi + 740. 113s.

Since Tinbergen published his now classical *Study of Instinct* in 1949 there has been a rapid development in ethological studies. For some time there has been an urgent need for a good textbook in this large and growing field. *Mechanisms of Animal Behavior* is a book which fills this gap very adequately, and will be a rich source of references and ideas for students for several years to come. It is significant that where Tinbergen's book was only 220 pages, the present volume runs to 740. Special praise should be given to the publishers for the excellence of the presentation: the print, typography and illustrations are some of the best of any textbook the reviewer has read. Fortunately the lucidity of the text and its matter both merit and equal this standard.

The book is inevitably long, in view of the breadth of its subject matter, but it nevertheless achieves a homogeneity of approach, which shows how far ethology has already progressed towards an intellectual synthesis in the science of behaviour. The synthesis has been achieved here not by including any lengthy theoretical chapters but by the

arrangement and selection of the material presented, so that the themes are part of its warp and woof. This is well illustrated by the fact that where most books would start with sensory and motor activities, and later progress to more complex aspects of behaviour, Marler and Hamilton consider first one of the central problems—the fact that animals are not passive, reflex machines, but vary in their sensitivity and responsiveness to external stimuli, and emit behaviour spontaneously. Variations in behaviour correlated with diurnal or circadian rhythms are first discussed, and then follow discussions of the more specific systems, reproduction, feeding, drinking and breathing, exploration and agonistic behaviour. The interaction of central and peripheral factors in the control of locomotion is used to illustrate how both aspects of the behaviour must be studied together.

The chapters on external stimuli fall into two sections, those dealing with releasing stimuli, under the heading of stimulus filtering, and those concerned with orientation. In all these chapters the correlation between structure and function is well illustrated. Marler's contributions in his own field of auditory communication are of the high standard one would expect.

The final section is concerned with the development of behaviour, dealing successively with its embryology, and the ontogeny of sensory and motor systems. This resolves the old question of separating the innate from the learned components of behaviour not by suggesting (as some ethologists have done) that the distinction is meaningless since it cannot be investigated experimentally, but by showing how fruitful it is to study the role of the environment during the development of the individual, particularly in early life.

Basically the ethological approach to behaviour is concerned with six problems—its immediate causal relationships, the underlying mechanisms, its genetic background, its development in the life history of the individual, its function (i.e. its survival value) and its evolutionary history. As Marler and Hamilton have shown these problems are not independent: the working of a sensory system (for example) is affected by its phylogenetic history, which imposes certain features upon its structure, and by its function, which imposes certain requirements upon it in terms of the physical nature of the stimuli to which it must respond. One might wish that the authors had dealt in more detail with mechanisms of evolution, social organization, and with the role of learning. Although the omissions were deliberate, and their inclusion would have lengthened the book considerably, a study at least of the ethology of learning (if only in the white rat) would have proved valuable, firstly by putting it in biological perspective and secondly by providing a path out of the maze of minitheories in which learning theorists now find themselves.

The omission of learning exemplifies what is admittedly a minor defect. The publishers' blurb suggests that the book represents a synthesis of ethology, psychology and physiology. Much relevant psychology has not been covered, and very little physiology. Indeed the selection of the physiological material included seems rather arbitrary. In the section dealing with feeding much behavioural data is presented, to which all people working in this area should certainly be introduced; but the section on brain mechanisms and hunger and thirst is brief, and is the sort of summary of Morgan's chapter that an undergraduate might produce. Similarly in the otherwise good section on auditory mechanisms there are about 12 pages on the structure of the vertebrate ear, but the electrophysiology of the auditory system receives only a brief and passing mention, even less than for insects. Although these defects lie in superficiality, in other cases a knowledge of necessary physiology is assumed, as, for example, in the section on hormones and behaviour, where the relationship between the hypothalamus and pituitary and gonadal hormones is nowhere specifically explained. To a reviewer for whom the word "mechanism" means something anatomical and physiological this must be a disappointment. But it would be churlish to allow this to detract from the very real merits of the book when it deals with actual behaviour, and what can be deduced from it.

D. M. VOWLES.

Animal Behaviour: A Synthesis of Ethology and Comparative Psychology. By Robert A. Hinde. London: McGraw-Hill. 1966. Pp. x + 534. 84s.

This book is a survey of "the area where psychology, physiology, and ethology overlap." The "synthesis" of the sub-title suggests system building, but it is more a collection of unconnected and cautious generalizations, established by critical discussion of a large number of well described experiments. There is little attempt at achieving coherence through theories, or through physiology; theories are occasionally introduced, and in

some detail (e.g. Deutsch's theory of learning), but for purposes of illustration (thus Deutsch's theory illustrates the possible complexity of goal seeking behaviour), rather than to provide the kind of close analysis and comparison between experiments that a theory, ideally, can achieve.

The obvious difficulty with such an empirical approach is to organize the facts in a palatable and illuminating way; perhaps such organization is impossible to sustain, in which case it is unfair to complain that interest begins to flag before reading through the full 446 pages. More relevant, since the book will be used largely for reference and as a textbook, is that it should contain adequate summaries and indexes. Good are the brief but informative summaries at the end of each chapter; though sometimes, particularly where results appear to conflict, it would be useful to have tables listing and briefly summarizing experiments. Not so good is the failure to indicate upon what page any particular reference occurs, so that, to see what Dr. Hinde says about a paper, one has to find out, from the author index, all the pages on which the author one is interested in is mentioned, and then go through them until one finds the reference (a small point perhaps, but it would seem simple to combine the list of references with the author index).

Like many non-theorists, the author sometimes seems over-preoccupied with conceptual matters. Thus, although he himself has shown how the old philosophical problems such as nativism vs. empiricism, the possibility of teleological explanation, etc., can be made to vanish by deft experiments which fully illuminate the area upon whose fogginess the problems were parasitic, he turns philosopher when he comes to the question of drives, and creates a straw man that Ryle himself would be proud of: "More serious still, however, is the temptation to look for the activating or energizing drive inside the organism. However willing the drive theorist may be to rally physiological data to his aid, to look inside the nervous system for drives, defined in terms of behaviour, is a logical mistake; and discussion about where drives originate is based upon a misconception. Although the drive theorist will not find any drives within the organism, he will find physiological correlates." True, and since he is doing so, such conceptual problems will be of interest only to philosophers.

Probably the question of most interest to experimental psychologists is the extent to which this book succeeds in bringing together their discipline with that of Ethology, and the mutual benefit, if any, obtained. There is nothing startling said about this, and little is made even of the obvious link provided by the current interests of some psychologists in comparative data as a means of testing models of behaviour. But by presenting side by side, and not too self-consciously, intelligent and detailed discussions of such topics as selective attention in humans and the "innate releasing mechanism" in birds, or skills and bird song, Dr. Hinde has written a book that will help to bring about a merging of the two disciplines in the minds of the next generation of students. A. W. STILL.

Manual of Psychophysiological Methods. Edited by P. H. Venables and Irene Martin. Amsterdam: North-Holland Publishing Co. 1967. Pp. ix + 557. £4 10s. od.

If psychophysiology is a science, psychophysiological methods ought to be scientific methods, and this book is not really about scientific method. Rather, it is about instrumentation for measuring those physiological quantities which psychologists use as physical indicators of a subject's psychological state. As is to be expected, the emphasis is on transduction and amplification, though there are chapters on data-processing and display, and telemetry. Phenomena considered are skin resistance and potential, heart rate, blood pressure, the plethysmogram, the electro-gastrogram, the electromyogram, the electro-oculogram and the electroencephalogram. The editors presumably have it in mind that readers will be psychologists with possibly small knowledge of physics and engineering, and have included an opening chapter called "Basic Physical Principles."

The contributors are mainly British. Some are essentially engineers, the others are biologically trained and are mainly psychologists. The engineers write well about engineering; the biologists are not always so successful. I am not in a position to comment on how sensibly engineers or biologists have written about biological matters.

Such is the pace at which technology advances that it is difficult enough for an engineer to keep up with new developments. To teach what he knows to a biologist so that the biologist can dispense with his (the engineer's) services is, except in rare cases, rapidly becoming impossible. I feel, therefore, that in so far as *Manual of Psychophysiological Methods* serves to introduce the subject to the young psychologist, to tell him what techniques are available and refer him to the literature, it is useful. But that in so far

as it seeks to teach him engineering, with the possible implication that he will then be able to do electrophysiology on his own, it is unfortunate. As apparatus becomes increasingly expensive, it becomes increasingly important to use it optimally. Not to do so may prove more expensive, in the long run, than to hire proper technical help.

P. E. K. DONALDSON.

The Senses. By Otto Lowenstein. London: Penguin Books. Pelican Original. 1966. Pp. 217. 5s.

When writing a paperback for the intelligent layman, or the young professional or school leaver, it always difficult to decide what to include, and how to present it without talking down. Professor Lowenstein makes it clear in the preface that this volume is not intended as a comprehensive elementary teaching textbook, but as a collection of selected essays in which some problems are discussed in more or less detail, to illustrate particular points. Thus he includes a long section in his own area of mechanoreception (56 pages) but only a short section on hearing (12 pages). The topics are presented from a standard zoological viewpoint relating structure to function. Some of the sections take the classical form of a dialogue between teacher and pupil, and these are unusually pleasantly and clearly written, having been based, one imagines, on Lowenstein's own teaching experience. The text is usually lucid, although condensation has sometimes led to errors. For example, the mosaic theory of visual acuity seems to be accepted to explain the minimum separable, while eye-movement is invoked to explain the minimum visible. In fact the scattering of light in the eye invalidates the first, while studies with the stabilized retinal image invalidate the latter. However, these are minor points possibly more relevant for an honours text. The book concludes with a philosophical dialogue on the nature of sensory experience. This is rather naive, but it is pleasant to find a zoologist who is at least prepared to discuss it.

D. M. VOWLES.

Human Spatial Orientation. By I. P. Howard and W. B. Templeton. London and New York: Wiley. 1966. Pp. 553. 84s. net.

This is a largely factual compilation and a very well-organized one. The first nine chapters cover basic psychophysiology, including visual localization, eye movements and visual direction, kinaesthesis and the vestibular system, and orientation with respect to gravity. These are followed by two chapters dealing with some wider aspects of geographical orientation (including certain of the disorders to which it is subject in consequence of brain injury) and two chapters dealing with the rather special problems of orientation and perceived shape. Finally, intersensory localization, the effects of displacement of the retinal image, and problems of orientation in the weightless state, are briefly considered.

To research workers in relevant fields, this book will be a gold-mine of well-selected experimental data, some nothing like as well known as they deserve to be. It also reflects a robust, yet perceptive, evaluation of what really is known about orientation in space, and of the techniques which have been devised to study it. And while primarily empirical, the authors show themselves by no means unfamiliar with current theory and prepared where necessary to offer shrewd criticism.

This book is in the best traditions of experimental psychology, as it evolved in nineteenth-century Germany, and is refreshingly free from the current obsession with theory and models. It is reassuring to know that two British psychologists can produce a book as good and solid as this one, even if one of them has alas since re-orientated himself in space.

O. L. ZANGWILL.

Attention: An Enduring Problem in Psychology. Edited by Paul Bakan. London: Van Nostrand. Insight Book No. 34. 1966. Pp. iv + 225. 14s. $1.75.

Each of the 10 sections in this book consists of a reprint (usually somewhat abridged) of a chapter or article from earlier works. Professor Bakan provides a short introduction to each section, and the book enables us to have a brief look at a number of the many approaches which various writers have made to the vast subject of attention. Each section provides topics (both historical and current) which deserve lengthy discussion but this is obviously impossible in a short review.

The first section is a good abridgement of William James's famous chapter, and the second a well selected part of Titchener's description of attention as "sensory clearness."

These historical introductions are followed by M. D. Vernon's rather comprehensive presidential address to Section J of the British Association (1959), and R. W. Gardner's (1961) discussion of "scanning principles" in relation to "field articulation." Then there is D. E. Broadbent's early paper (1953) on the relation of attention to classical conditioning.

These are followed by three sections dealing more specifically with the relation of attention to individual differences. I Maltzman and D. C. Raskin (1965) discuss the "orienting reflex," Bakan, J. A. Belton and J. C. Toth (1963) discuss Extraversion and Introversion, and J. Silverman (1964) discusses Schizophrenia.

Finally there is a section by Hernández-Péon (1964) summarizing some of the physiological work of his colleagues and himself up to that date; and the book concludes with a section by J. A. and D. Deutsch (1963) in which they discuss Broadbent's "filter" theory, the work of Hernández-Péon's school, and also other recent work.

Professor Bakan is to be congratulated on producing so good a book on such a variety of topics in such a small space.

G. C. Grindley.

Amnesia. Edited by C. W. M. Whitty and O. L. Zangwill. London: Butterworths. 1967. Pp. x + 217. 64s.

This authoritatively informative book fills a distinct gap in the literature about human memory. Clinical and pathological aspects of memory have long posed problems for physicians, neurologists, psychiatrists, lawyers, and psychologists. From practical diagnostic and medico-legal points of view, these problems are urgent in their own right. From a more scientific viewpoint, they are both valuable and difficult: valuable because they force attention on the full intricacies of memory and memory losses as they occur in the total functioning of the person and, so, act as corrective to the sometimes myopic theories fostered by purely theoretical speculation or laboratory experiment: difficult because they demand the exercise of sensitive naturalistic observations and clinical probings under conditions which are not always conducive to the drawing of firm conclusions. Over the years, these valuable clinical studies have cumulated and shown an increasing semblance of order; more recently, these studies have been augmented by relevant sophisticated work in neurology and experimental psychology. The result has been a growth of knowledge about amnesia which has practical applications, and also important implications for views of normal psychological functioning. There has, however, been no single book which broadly surveys this contemporary knowledge. *Amnesia* is just such a book. Its nine chapters have been written by a total of ten specialists, all familiar at first-hand with their respective topics. These authors have been coaxed into complimenting each other so as to provide a readable, comprehensive, up-to-date account of knowledge about abnormal memory losses. This account is an invaluable reference for those who must deal with practical problems of amnesia in hospital or court-room. It is a richly convenient source of clinical and experimental data for those concerned with scientific problems of memory. And it will surely stimulate new thinking and enquiry in this entire area which has such relevance for central questions about psychological functioning. For all of these reasons, this book is timely and eminently worthwhile.

Throughout the book, memory is depicted as a biological capacity, of very considerable complexity, which is related to the physical structure and function of the central nervous system. From the different chapters, there emerges a group of recurring themes and also an overview of memory organization which, though tentative and open to future elaboration, is faithful to a wide variety of cited findings and current ways of thinking about functional issues. So much of the book is itself a review that it would be inappropriate, here, to do more than indicate its contents. The opening chapter, by L. Weiskrantz, considers experimental studies of amnesia in animals, mostly monkeys; it cites the extent to which various amnesic conditions have been selectively produced by experimental ablations; and it concludes with a rough model of memory which involves notions of short-term traces, long-term traces and their autonomous strengthening, and "noise" levels. These notions are echoed in subsequent chapters which are more exclusively concerned with human findings. Explicit consideration of human amnesia begins with an overview, by C. W. M. Whitty and W. A. Lishman, of the transient and persistent memory losses accompanying cerebral diseases of various kinds: and it continues with more detailed considerations by O. L. Zangwill on the amnesic (Korsakoff) syndrome; by Whitty and Zangwill on traumatic amnesias associated with head injuries; by Brenda Milner on amnesia following operation on the temporal lobes; and by Moyra Williams on memory

disorders associated with electroconvulsive therapy. J. B. Brierley then turns attention to the evidence obtained from neuropathological examination of the human brain in cases of amnesia; E. Stengel attends, more briefly, to psychogenic loss of memory; and, finally, medico-legal aspects of amnesia are reviewed by T. C. N. Gibbens and J. E. Hall Williams. Throughout, the editors have kept tight control of their distinguished contributors so as to avoid the disharmonies of multiple authorship and produce a work which is both coherent and remarkable for the sheer amount of useful information and suggestion it contains. The editors have done a great service by bringing so much material together and so well. This review can best conclude by endorsing the words of Sir Charles Symonds who, in his preface, writes "I am sure that this book will gain the success it deserves and I hope that it will reappear in further editions, to keep pace with the advance of knowledge in its field."

I. M. L. HUNTER.

Brain Function. Volume III: *Speech, Language and Communication*. Edited by Edward C. Carterette, UCLA Forum in Medical Sciences. No. 4. University of California Press. London: Cambridge University Press. 1967. Pp. xiii + 279. 96s.

This is the proceedings of yet another "inter-disciplinary" conference, held in November, 1963. The disciplines represented comprise neurology, psychiatry, communications engineering, neurophysiology, phonetics, linguistics, experimental psychology and biophysics. Communication between their representatives seems on the whole very much better than might have been feared.

To experimental psychologists, the most interesting chapters will probably be those by Ira Hirsch on audition in relation to speech perception and Eric Lenneberg on developmental factors in speech perception. But for those with more esoteric interests may be mentioned Norman Geshwind's sympathetic account of Carl Wernicke and his work and Roman Jakobson's discussion of linguistic types of aphasia. There are also chapters by C. E. Osgood, C. A. Ferguson and S. M. Lamb.

O. L. ZANGWILL.

Annual Review of Psychology, Volume 18. Edited by P. R. Farnsworth, O. McNemar and Q. McNemar. Palo Alto, California: Annual Reviews Inc. 1967. Pp. 606. $9.00 ($8.50 in U.S.A.).

This volume contains the following chapters: Perception, by D. E. Johannsen; Visual Sensitivity, by F. A. Mote; Comparative Psychology and Ethology, by J. P. Scott; Developmental Psychology, by H. W. Stevenson; Educational Psychology, by R. C. Anderson; Projective Methodologies, by S. Fisher; Personnel Selection, by R. M. Guion; Test Theory, by J. A. Keats; Value: Behavioral Decision Theory, by G. M. Becker and C. G. McClintock; Group Dynamics, by H. B. Gerard and N. Miller; Psychotherapy, by D. H. Ford and H. B. Urban; Classification of the Behavior Disorders, by J. Zubin; Audition, by J. J. Zwislocki; Organizational Psychology, by R. P. Quinn and R. L. Kahn; and Personality, by G. S. Klein, H. L. Barr and D. L. Wolitzky. Hardly more than five of these deal with areas of psychology that are often represented in this *Journal*, and the emphasis is strongly on what one might call professional psychology. In the face of all this one feels like the unqualified in pursuit of the unreviewable.

The reviewer's uneasiness over his task is equalled by the authors' over theirs, and indeed most of the things one finds oneself wanting to say about the *Annual Review* have at least as much to do with the task the authors were set as with the way they have tackled it. In the first place there is the selection of topics. Some of them—Perception, or Developmental Psychology—are so ill-defined and could be understood as including work so heterogeneous that it is difficult even to imagine what a satisfactory review would be like. As a result some of the authors are forced on to the defensive throughout. Their chapters begin with nervous introductions in which they excuse themselves for not covering the area satisfactorily, and end with such apologetic statements as Johannsen's: "A great deal of work has been done, much of it sophisticated and critical, but it does not seem to hang together in any very systematic way." One sympathizes: what else could she possibly have said about perception?

In the second place, although the policy of the *Annual Review* seems to be that authors should by and large confine themselves to reviewing material that has appeared in the period since their topic was last dealt with, it by no means always follows that the period has any special significance for the development of that topic. Few authors were able to find any distinctive features in the period under review. Becker and McClintock are

something of an exception, but Decision Theory has grown rapidly since 1961. It is more common for authors to be forced to this sort of conclusion: "Research on personality and social development is characterized this year by extensive and detailed studies of traditional topics, rather than by the appearance of anything that resembles a new look." This sounds like a dispatch from the Paris fashion shows in a year when the couturiers are on strike.

Thirdly, authors are faced with deciding how much space to allow themselves for commenting on the material they review. Clearly, for many of the topics chosen, an author cannot be expected both to refer to most of the work in his area and to offer any estimate of its value. Many of the differences between chapters arise from different attempts to solve this problem. Some authors (Stevenson, for example) see their task as being to summarize as much material as space allows them. One wonders whether the reader would not have gained at least as much from an afternoon in the library glancing through the paper summaries in relevant journals. Others deliberately restrict the scope of their chapters to leave themselves more time for discussion; thus, for Scott Ethology becomes work on the relationship between behaviour and evolution, and for Anderson Educational Psychology is mainly programmed learning. In Scott's chapter the restriction seems to me to produce an eccentric result; amongst other things it leads him, rather oddly, to criticize Lorenz's hydraulic model for having "no counterpart in the nervous system of any known animal." In Anderson's it is worthwhile since it gives him the opportunity for quite detailed discussion of a number of topics in the field of programmed learning. Perhaps the worst solution to this problem is that of Gerard and Miller who state their intention only to discuss work that they consider to be of high quality but then leave themselves no space for saying why they think the work they choose is good. There is no point in simply stating that a paper is "fine" or "interesting" unless one has time to say why.

There is no sign that previous criticisms of the style of most of the writing in the *Annual Review* have had any effect. When an author writes, "There is very little research investigating the possible instrumental value of conformity as an illicit form of eliciting a positive self-evaluation from someone else," he must expect to lose his reader's attention. It is not pedantry that makes one object to sentences like this, but a concern that language should continue to be useful.

M. H. SHELDON.

Contemporary Approaches to Psychology. Edited by H. Helson and W. Bevan. Princeton, New Jersey and London: Van Nostrand Company. 1967. Pp. xii + 596. £5 16s.

This is the successor to *Theoretical Foundations of Psychology* (1951) and contains 12 chapters aimed at the advanced student. Two of the contributions come from the editors ("Perception" by Helson; "Behavior in Unusual Environments" by Bevan) and the remainder range from "Engineering Psychology" by Adams to Guilford's chapter on Creativity.

Unfortunately, the approach of these contributions is not quite "contemporary." Collections of essays seem to have the habit of taking a long time to get to press, particularly when the idea for writing them comes from a publisher rather than the authors. Much of the work here has been written 4 years ago or more, and the chapters that attempt a review of the current literature are inexcusably out of date. Several otherwise excellent contributions have suffered badly from the long delay in publication, amongst them the chapter on Neurophysiology by William and Martha Wilson, and "Motivation and Affectivity" by Hall. Both would have been useful reviews if published a few years ago, but an idea of the disadvantages from which they now suffer may be gained from the fact that the Wilsons' discussion of self-stimulation mentions none of Deutsch's work, while the section on the physiology of learning has no references later than 1962.

The title of the book is a bit misleading for another reason. A number of the chapters are not concerned with contemporary approaches at all, and settle either for a rather general discussion of long-standing problems or for a straightforward historical survey. Bitterman's entertaining account of "Learning in Animals" is of the last kind, and can be recommended equally to those wishing an introduction to the classical problems of learning theory, and to those contemporary anti-theorists who seem to need reminding that the problems are still with us. Another line of approach that can be seen in the various chapters is the "Methodological," in which an attempt is made to define the proper subject matter and methods of a particular subject. For example, Phares in "The Deviant Personality" is worried mainly by the old problem of defining abnormal behaviour; he

has read his Szasz and gives the reader a lot of far from innocent fun by coolly observing that the concept of mental illness is a "hoax." It is supposed to be healthy for theorists to look at their underlying assumptions in this kind of way, but it is easy for the self-examination to become a bit neurotic. If they share this reviewer's rather unsophisticated enthusiasm for mere experimental detail, readers may well feel cheated by the following illuminations:

"To be part of Science, the results of observation must be formulated in terms that will place them in the public domain" (p. 91 "Behavior Genetics": Hirsch and Erlenmeyer-Kimling).

"Behavior is always, in an essential sense, a physical phenomenon, for it is always instigated, immediately or otherwise, by a physical energy change impinging on the organism" (p. 386, "Behaviour in Unusual Environments": Bevan).

"Science is an intellectual discipline" (p. 181, "Verbal Learning": Bower).

The trouble with these heart-searchings is that they are simply used to pad out an Introduction, and are then put aside. Lanzetta and Sieber in their chapter on Social Psychology ask "What is an Experiment?" (p. 529), but have forgotten their answer by the time it comes to describing as an experiment something that tests the "strain towards symmetry" model of social behaviour by showing that "if X and Y like one another, they will tend to similarly evaluate an event which they have just discussed" (p. 546). If there are any genuine points to be made about procedural difficulties in any subject, there is no reason why they should not be illustrated by examples all the way through the chapter, as they are in Garner and Creelman's discussion of "Problems and Methods of Psychological Scaling." An example of what can be achieved in this way is Hall's account of the puzzling way in which one-group versus two-group experiments can give different answers to a variety of questions about secondary reinforcement, such as the relation of frequency of primary reinforcement to the strength of a secondary reinforcer.

The book has a large number of well produced illustrations, including a few photographs, but large errors in the text have been allowed to get through. An omission on p. 45 makes a paragraph incomprehensible, and the attempt on p. 549 to illustrate the meaning of cognitive dissonance gets into such a mess that it is even more hilarious than these examples usually are.

M. J. MORGAN.

The Hypnotic Investigation of Dreams. By C. Scott Moss. London and New York: Wiley. 1967. Pp. xi + 290. 60s.

In the first part of this book, the author reviews both clinical and experimental studies of hypnosis and dreams, with special reference to hypnotically induced dreams. His treatment, though fairly standard, is full and surprisingly critical. The second part contains a dozen papers reprinted from the Journals which have to do with topics discussed in the preceding review. These are rather uneven, though one or two are worthwhile. One thing that rather surprised this reviewer was the absence of any mention of Morton Prince, whom he would regard as the Founding Father of the Hypnotic Investigation of Dreams. It is odd that this once so celebrated man is now so completely forgotten. But perhaps Miss Beauchamp may yet appear in paperback.

O. L. ZANGWILL.

The Dynamics of Behavior Development: An Epigenetic View. By Zing-Yang Kuo. Random House Studies in Psychology (Consulting Editor, L. J. Stone) New York: Random House. 1967. Pp. xii + 240. $2.45.

Dr. Kuo was a convert of Behaviourism who became especially interested in problems of individual development. Although he is remembered particularly for his experiments on bringing up rats and cats together (which dished McDougall and delighted pacifists), Dr. Kuo also published a number of studies of embryonic behaviour, especially in birds. But since the outbreak of the last war Dr. Kuo appeared to have vanished, and as he himself says in his Preface, many of his friends have wondered what had become of him and younger students of behaviour may not even recognize his name. It transpires, however, that for the past 20 years Dr. Kuo has "resided as an uninvited guest in the British Colony of Hong Kong" where he has been "enjoying the rare luxury of the freedom of silence."

Readers of this thoughtful little book will find that the author has indulged in this freedom to considerable advantage. He has clearly left the crudities and over-simplifications of Watson's Behaviourism far behind him, and shows an excellent understanding of modern attitudes. Indeed he now professes an "organismic" philosophy and places great

emphasis on the need for an inter-disciplinary approach to the problems of behaviour. It would be nice if this "uninvited guest" whom we have unknowingly harboured so long could be cordially invited to attend the XIX International Congress of Psychology in London.

O. L. ZANGWILL.

Self-Evaluation: Concepts and Studies. By James C. Diggory. London and New York: Wiley. 1966. Pp. xiii + 477. 80s.

Diggory and his associates have carried out a series of studies in which phenomena connected with the self are experimentally manipulated. Their main dependent variable is estimated probability of success at laboratory tasks—rather like level of aspiration. Studies are reported of the effects of success and failure, generalization from one task to another, the effects of observing the success or failure of others, and of mental patients. The experiments are carefully conducted, but have not used some of the more interesting ideas coming out of clinical studies of the self by such psychologists as Maslow and Erikson.

MICHAEL ARGYLE.

The Psychology of Interpersonal Behaviour. By Michael Argyle. Harmondsworth: Penguin Books Ltd. 1967. Pp. 223. 4s. 6d.

Anyone taking a casual glance at this book could easily underestimate its importance. At first sight it is largely a catalogue of various ways in which people have social contact with one another and of research which has been done on the problems which arise during such contacts. It is, however, much more than this: it represents the first full survey of the leading ideas and results of a programme of research into *social skills* started at Oxford by the author and Dr. E. R. F. W. Crossman with initial backing from the Human Sciences Committee of the former Department of Scientific and Industrial Research and now from the Social Sciences Research Council. The aim of this research has been to treat social contacts as a type of situation in which social skills are exercised, and to analyse these in the terms which have been developed over the past quarter century for sensori-motor skills practised by individuals in laboratory experiments, in using service equipment and in industry. By doing so, it brings a welcome breath of fresh air to the study of social behaviour and a closer link between this and the main body of psychology than has existed hitherto.

After a brief editorial foreword and a preface, the text is divided into ten chapters. These are followed by an "epilogue," valuable lists of references and suggestions for further reading, and short indices of names and subjects. The key to the book is in Chapters 5 and 6. In the former, basic ideas about social skills are set out, and individual and social skills are compared under six main heads:

(1) The need for defined goals or aims—skill is regarded as essentially a flexible means of attaining goals.
(2) The development of selectivity in perception.
(3) "Translation processes" relating perception to action.
(4) Actual motor responses.
(5) Feedback and the corrective effects it produces.
(6) The timing of action within a complex overall pattern.

The author then goes on to list certain needs which he suggests are special features of social skills and not present in manual skills. They are to (*a*) establish rapport, (*b*) keep the other person "in play," (*c*) motivate him, (*d*) reduce anxiety and defensiveness, and (*e*) be concerned about the impression made on the client. The chapter ends with an outline of what the author regards as marks of social competence: perceptual sensitivity, warmth and rapport, a repertoire of social techniques, flexibility, energy and initiative, and "smooth response patterns."

Chapter 6 describes some of the extremely interesting experiments which have been carried out by the author and his team. The other chapters deal mainly with the work of others, but clearly aim at re-casting it in a way which enables it to be viewed in terms of social skills.

The book shows some evidence of haste. For example, statements are occasionally a little sweeping and it is not always made clear that the differences between groups or effects being discussed are statistical rather than absolute. Again, many of the "special" features of social skills look on reflection as if they could be subsumed under those common

to motor skills as well. These faults are trivial, however, compared with the substantial re-thinking of almost the whole field of social psychology from what is potentially a very revealing angle.

A. T. WELFORD.

The Causes of Behaviour II. Second edition. Edited by Judy F. Rosenblith and Wesley Allinsmith. Boston: Allyn and Bacon. 1966. Pp. xv + 608. $6.95.

This is a large book, and a solid one. Like so many other present-day psychological publications, its main content consists of a collection of papers and extracts by different authors. The intended readers are students in developmental, educational and child and adolescent psychology. The book offers them information ranging in date from 1905 to 1965, and in topic from L. S. Kubie's Introduction to "An Application of Psychoanalysis to Education" (by R. M. Jones) to D. G. Freedman's paper on "Constitutional and Environmental Interactions in Rearing of Four Breeds of Dogs." Great names rub shoulders with the less eminent, opinion and theory are intermingled with experimental studies, and while some selections run to 14 pages, others occupy less than a page. Clearly a book of this kind must depend for its efficacy on the organization achieved by the editors and the guidelines they provide for the reader. These are largely, but not uniformly, successful here. Each section has a useful introduction in which comments of a general kind are made on the field of study of the coming selections, and their content is briefly discussed. In addition to providing this context for reading, the introductions are taken as an opportunity to define terms and to mention supplementary reading which will extend the student's knowledge of the field.

Four reflective articles introduce the book as a whole. They are by Allport ("Psychological Models for Guidance"), Skinner ("Behaviorism at Fifty"), Bruner ("Freud and the Image of Man"), and Rogers ("Toward a Modern Approach to Values: the Valuing Process in the Mature Person"), and have been chosen by the editors to emphasize the diversity of approach which is to follow. The subsequent organization is in terms of eight categories of behavioural determinant: Biological Bases, Learning, Interpersonal Experiences, Settings and Specific Stimuli, Group Memberships, Age or Developmental Stage, Sex, and Intelligence. After a section discussing "Motivational Resultants," the book closes with twelve selections which deal with specific educational implications. The student of educational psychology would be well advised to make these last selections his start on a book in which the editors "have attempted to offer a collection that changes pace frequently and that ranges widely in topics covered, in techniques of investigation used, in levels of difficulty, and in the views espoused by the authors."

There are occasions when this intention leads to unsatisfactory results. Some of the sections are weakened in their effect by the gulf in approach or content between their constituent selections. Although the introductory overviews are offered to help the reader to integrate the material, their purpose is defeated when the scatter has the extent of, for example, the selections grouped under "Settings and Specific Stimuli as Determinants of Behaviour." Here the group psychology of the class, television, comic books, psychoanalysis, and cultural deprivation are lumped together in a glorious hotchpotch, and all within the space of 30 pages. The effect is comparable to the frustration experienced if one attempts to botanize from the window of a fast-moving train. In short, what is sometimes achieved is not representative variety but fragmentation. Perhaps the "Readings" format is inappropriate for the scope of such a wide subject as that of this book. Excellence of individual selections is not a substitute for the reader's grasp of theme, and the most useful collections now on the market are those where content is much more delimited by topic or recency. Perhaps the maligned textbook still remains a more suitable vehicle for the presentation of the view in breadth. Yet the selections in this second edition of Rosenblith and Allinsmith have been largely determined by consumer reaction. The editors were guided by the evaluations of their own students, by the answers given to a questionnaire by academic colleagues from other Universities, and by professorial response to the publishers' request for reactions to the book. It must be presumed, then, that the psychological thesaurus compiled of discrete papers satisfies a real need—amongst American students at any rate.

D. J. BRUCE.

PUBLICATIONS RECEIVED

BOOKS

Theories of Child Development. By Alfred L. Baldwin. London and New York: Wiley. 1967. Pp. xii + 618. 70s.

The Paradox of Guilt: A Christian Study of the Relief of Self-Hatred. By Malcolm France. Foreword by Frank Lane. London: Hodder & Stoughton. 1967. Pp. 128. 25s.

The Emotional Health of Physicians. By John C. Duffy and Edward M. Litin. Springfield, Illinois: Thomas. 1967. Pp. ix + 88. $5.50.

Occupational Information for the Mentally Retarded: Selected Readings. Edited by Lotar V. Stahlecker. Springfield, Illinois: Thomas. 1967. Pp. xxii + 816. $42.00.

The Play Theory of Mass Communication. By William Stephenson. London: University of Chicago Press. 1967. Pp. x + 225. 37s. 6d.

Approaches to Psychopathology. Edited by James D. Page. London: Columbia University Press. 1967. Pp. xii + 304. 56s.

Vectoriasis Praecox or The Group of Schizophrenias. By Benjamin B. Wolman. Springfield, Illinois: Thomas. 1966. Pp. xi + 371. $12.50.

Introduction to Statistical Analysis and Inference for Psychology and Education. By Sidney J. Armore. London and New York: Wiley. 1967. Pp. xx + 546. 72s.

Medieval Minds: Mental Health in the Middle Ages. By Thomas F. Graham. London: Allen & Unwin. 1967. Pp. 112. 30s.

The Roots of the Ego. By Carl Frankenstein. Baltimore: Williams & Wilkins (Edinburgh: Livingstone). 1966. Pp. xi + 286. 62s.

Eclectic Psychiatry. By Rudolf Kaelbking and Ralph M. Patterson. Springfield, Illinois: Thomas. 1966. Pp. xxv + 891. $17.50.

Educational Psychology in the Classroom. 3rd Edition. By Henry Clay Lindgren. London and New York: Wiley. 1967. Pp. xx + 686. 68s.

Hypnosis: A Clinical Study. By S. Sunder Das. London: Asia Publishing House. 1967. Pp. 155. 25s.

Korperbau und Charakter: Untersuchungen zum Konstitutionsproblem und zur Lehre von den Temperamenten. By E. Kretschmer. Berlin: Springer-Verlag. 1967. Pp. xvi + 484. DM. 48. $12.00.

Invertebrate Zoology. By Paul A. Meglitsch. London: Oxford University Press. 1967. Pp. xx + 961. 88s.

Methods for Experimental Social Innovation. By George W. Fairweather. London and New York: Wiley. 1967. Pp. x + 250. 60s.

Psychology of Adjustment. 2nd Edition. By James M. Sawrey and Charles W. Telford. Boston: Allyn & Bacon. 1967. Pp. x + 466.

The Neural Basis of Behaviour. By Lloyd S. Woodburne. London: Prentice/Hall International. Merrills International Series. 1967. Pp. vi + 378. 94s.

The Psycho-Analytical Process. By Donald Meltzer. London: Heinemann Medical Books. 1967. Pp. xvii + 109. 20s.

Pharmacotherapy of Depression. Edited by Jonathan O. Cole and J. R. Wittenborn. Springfield, Illinois: Thomas. 1966. Pp. xi + 189. $8.50.

Current Research in Motivation. Edited by Ralph Norman Haber. London and New York: Holt, Rinehart and Winston. 1966. Pp. xii + 800. £5.

The Nature of Perceptual Adaptation. By Irvin Rock. London: Basic Books. 1967. Pp. xii + 289.

Biological Foundations of Language. By Eric H. Lenneberg. With Appendices by Noam Chomsky and Otto Marx. London and New York: Wiley. 1967. Pp. xvi + 489. 120s.

PAPERBACKS

Nerve, Muscle and Synapse. By Bernard Katz. London: McGraw-Hill. 1966. Pp. xi + 193. 18s. soft cover. 40s. hard cover.

Insearch: Psychology and Religion. By James Hillman. London: Hodder & Stoughton. 1967. Pp. 117.

New Directions in Psychology III. By George Mandler, Paul Mussen, Nathan Kogan and Michael A. Wallach. New York and London: Holt, Rinehart & Winston. 1967. Pp. ix + 289. $3.95.

Acquisition du Langage et Développement de la Pensée: Sous-systèmes Linguistiques et Opérations Concrètes. By H. Sinclair de Zwart. Paris: Dunod. 1967. Pp. vi + 168. 23F.

Psychologie et epistemologie genetiques Themes Piagetiens. Hommage à Jean Piaget avec une bibliographie complet de ses oeuvres. Paris: Dunod. 1966. Pp. xxi + 421. 39F.

Readings in the Psychology of Parent-Child Relations. By Gene R. Medinnus. London and New York: Wiley. 1967. Pp. xiii + 371. 35s.

Psychoanalysis and Culture. Edited by George B. Wilbur and Warner Meunsterberger. London and New York: Wiley. Science edition. 1967. Pp. xii + 462. 21s.

Proceedings of the 1966 Invitational Conference on Testing Problems. Princeton, N.J.: Educational Testing Service. 1967. Pp. vii + 123. $1.50.

The Psychology of Interpersonal Relations. By Fritz Heider. London and New York: Wiley. 1967. Pp. ix + 322. 16s.

Insight Books. London: D. Van Nostrand. 1967.

- No. 37. *Language and Thought.* Edited by Donald C. Hildum. Pp. vi + 200. $1.95. 16s.
- No. 38. *Psychological Perspectives on the Person.* By Norman Tallent. Pp. x + 293. $2.95. 24s.
- No. 39. *Hormones and Behavior.* Edited by Richard E. Whalen. Pp. vi + 266. $2.95. 24s.
- No. 40. *Creativity and Conformity.* By Clark Moustakas. Pp. xiii + 142. $1.95. 16s.
- No. 36 *Psychological Needs and Cultural Systems: A Case Study.* By Joel Aronoff. Pp. xiii + 241. $2.95. 24s.

The Analysis of Fantasy. By William E. Henry. London and New York: Wiley. Science Editions. xiii + 277. $2.45. 192.

Mental Health Book Review Index: An Annual Bibliography of Books and Book Reviews in the Behavioral Sciences. Vol. 12, 1967. With an editorial: Bibliographic Foundations for an Emergent History of the Behavioral Sciences. Pp. xxv+89. $8.00. Published by the Council on Research in Bibliography, Inc., c/o Research Center for Mental Health, New York University, New York, N.Y. 10003. Sponsored by the World Federation for Mental Health, International Council of Psychologists, and Research Center for Mental Health, New York University. MENTAL HEALTH BOOK REVIEW INDEX, c/o Paul Klapper Library, Queens College of the City University of New York, Flushing, New York 11367, U.S.A.

MONOGRAPHS

Habituation to complex vestibular stimulation in man: transfer and retention of effects from twelve days of rotation at 10 RPM. By Fred E. Guedry. *Perceptual and Motor Skills, Monogr. Suppl.* 1-V21, 1965. Pp. 23. $1.00

Mnemonic organization as a determinant of error-gradients in visual pattern perception. By E. Rae Harcum. *Perceptual and Motor Skills, Monogr. Suppl.* 5-V22, 1966. Pp. 24. $1.00.

Reinforcement-test sequences in paired-associated learning. By Chizuko Izawa. *Psychol. Reports. Monogr. Suppl.* 3-V18, 1966. Pp. 40. $1.50.

Human Performance and Behaviour in Hyperbaric Environments. By John Adolfson. Acta Psychologica Gothoburgensia VI. Stockholm: Almqvist & Wiksell. 1967. Pp. 74. Sw. Kr. 28.0.

INDEX TO AUTHORS

.n
/il